PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/70, A61K 39/29, C07K 14/18 | A1

(11) International Publication Number: WO 95/11998

(43) International Publication Date: 4 May 1995 (04.05.95)

(21) International Application Number: PCT/US94/12268

(22) International Filing Date: 26 October 1994 (26.10.94)

(30) Priority Data:
08/143,412 26 October 1993 (26.10.93) US

(71) Applicant: UNITED BIOMEDICAL, INC. [US/US]; 25 Davids Drive, Hauppauge, NY 11788 (US).

(72) Inventors: WANG, Chang, Yi; 47 Snake Hill Road, Cold Spring Harbor, NY 11742 (US). ZAMB, Timothy, J.; 54 Bonnie Lane, Stony Brook, NY 11790 (US). YE, John; 29-02 200th Street, Bayside, NY 11360 (US). KAMINSKY, Stephen, M.; 5355 Henry Hudson Parkway #12E, Bronx, NY 10471 (US). HOSEIN, Barbara; Apartment 6A, 196 E. 75th Street, New York, NY 10021 (US). NIXON, Douglas, F.; 112 Bayview Avenue, Northport, NY 11768 (US). KOFF, C., Wayne; 7 Night Herron Drive, Stony Brook, NY 11790 (US). KOWALSKI, Jacek; 25 Ashley Circle, Commack, NY 11725 (US). WALFIELD, Alan, M.; 45 Schiller Avenue, Huntington Station, NY 11746 (US).

(74) Agent: LIN, Maria, C., H.; Morgan & Finnegan, 345 Park Avenue, New York, NY 10154 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.

(54) Title: STRUCTURED SYNTHETIC ANTIGEN LIBRARIES AS DIAGNOSTICS, VACCINES AND THERAPEUTICS

(57) Abstract

The present invention relates to "structured synthetic antigen libraries" (SSAL) composed of related peptides synthesized simultaneously in a single peptide synthesis. This "structured" library contrasts to those libraries previously described as "random peptide libraries" in that the order or structure within a synthetic antigen is provided by invariant amino acid residues that define the framework sequence of the synthetic antigen. The specific amino acids and their frequency of appearance at a variant locus within aligned peptide sequences is defined by the primary sequences of the several variants that make up the alignment used to construct the antigen peptide library. A method of constructing an open diagnostic, vaccine or therapeutic for a mutational infectious agent is also provided. The invention further provides the SSAL in diagnostic methods, kits, vaccination methods, vaccine compositions and pharmaceutical compositions. The libraries are prepared from variable domains in proteins and provide improved vaccines, diagnostics and therapeutics for infectious agents, etc., from such proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

STRUCTURED SYNTHETIC ANTIGEN LIBRARIES AS DIAGNOSTICS, VACCINES AND THERAPEUTICS

FIELD OF THE INVENTION

The present invention relates to "structured synthetic antigen libraries" (SSAL) composed of related peptides synthesized simultaneously in a single peptide synthesis. This "structured" library contrasts to those libraries previously described as "random peptide libraries" in that the order or structure within a synthetic antigen is provided by invariant amino acid residues that define the framework sequence of the synthetic antigen. The specific amino acids and their frequency of appearance at a variant locus within aligned peptide sequences is defined by the primary sequences of the several variants that make up the alignment used to construct the antigen peptide library. A method of constructing an open diagnostic, vaccine or therapeutic for a mutational infectious agent is also provided. The invention further provides the SSAL in diagnostic methods, kits, vaccination methods, vaccine compositions and pharmaceutical compositions. The libraries are prepared from variable domains in proteins and provide improved vaccines, diagnostics and therapeutics for infectious agents, etc., from such proteins.

BACKGROUND OF THE INVENTION

Synthetic peptides have been used increasingly to map antigenic or immunogenic sites on the surface of proteins. These approaches have used: (1) nested sets of peptides as in "site-directed serology" to identify diagnostically useful epitopes; (2) overlapping sets of peptides to scan the length of a protein antigen to locate antigenic sites (a technique colloquially known as pepscan), and (3) totally random peptide libraries which can be systematically analyzed to characterize particular antigenic sites.

In this regard, synthetic peptides provide a powerful tool for the diagnostic field, particularly in the areas of diagnostic virology, blood screening and others. For example, using site-directed serology, sensitive and specific tests have been developed for the envelope and core proteins of human immunodeficiency virus (HIV) [see, e.g., U.S. Pat. Nos. 4,735,896 and 4,879,212], for the gag protein of human T-lymphotropic viruses I and II (HTLV I/II) [see, e.g., U.S. Pat. No. 4,833,071], and the structural and non-structural proteins of the hepatitis C virus (HCV) [see, e.g., U.S. Pat. No. 5,106,726].

In certain cases, peptide-based assays are unmatched in their ability to differentiate between or among closely related viruses, a problem often associated with other common diagnostic tests based on either viral lysates or recombinantly-produced proteins. While having many advantages, there are still pitfalls to peptide-based immunodiagnostics, as well as the other formats for immunodiagnostic tests, when dealing with infectious agents that vary significantly over time or may have many serotypes. Peptide-based and other immunodiagnostics are static by design and can not accommodate such variation.

As is well-known, many proteins from infectious agents, or sites on those proteins, can vary in sequence by strain. At a given time or in a particular geographic locale, a particular antigen may contain one or multiple point mutations relative to an arbitrary prototype strain. Dynamic variation thereby creates an extremely complex antigenic profile for a given site for which sensitive or specific detection may be increasingly difficult as the site "drifts" further from the prototype. Nonetheless, each antigen maintains a "fundamental motif" or "structural framework" that is representative of the prototype sequence from which the variants arise. The structural framework

consists of generally invariant residues which can be exploited in conjunction with the inherently variable residues to design improved diagnostic tests that overcome the limitations of the present tests.

Accordingly, the present invention provides a means to construct a synthetic peptide library termed a "structured synthetic antigen library" (SSAL) which addresses the need for "open" diagnostics, i.e. diagnostic kits which can expressly accommodate antigenic variation.

Site-directed synthetic peptides also find use as vaccines. Vaccine development and delivery has been highly successful as a public health strategy in the control of acute viral and bacterial infectious diseases. However, several major globally important infectious diseases has proven elusive to attempts at designing effective vaccines. A recurrent theme among certain pathogens, contributing to unsuccessful or limited success of vaccines, is antigenic variation. Extensive antigenic variation, for example, is a hallmark for example of HIV (AIDS), rhinovirus (the common cold), influenza virus (flu), <u>Borrelia</u> <u>burgdorferi</u> (Lyme disease), <u>Chlamydia</u> (pelvic inflammatory disease), <u>Plasmodium</u> <u>falciparum</u> (malaria), <u>Trypanosoma</u> <u>cruzi</u> (Chagas' disease). While biotechnology may enable production of large quantities of purified antigens containing protective epitopes, antigenic variation has thus far proved to be a significant obstacle for vaccine development in certain pathogens.

Effective vaccines generally must induce long term memory responses so that a subsequent protective response is elicited immediately upon exposure to a pathogen. A single recombinant protein or a single peptide confers only limited protective immunity to pathogens that exhibit extensive antigenic variation because of their limited capacity to induce heterotypic immune responses. Synthetic peptide

technology using SSALs offers a powerful tool to develop the next generation vaccines because they encompass a broad range of sequences required to overcome problems associated with antigenic variation as previously discussed for diagnostic markers.

Therefore, a novel strategy for vaccine-induced protection against pathogens having multiple strains, forms, serovars, serotypes, etc., is required. The "structured synthetic antigen library" (SSAL) provides a solution to the problem of developing vaccines for such pathogens for the following reasons. First, SSALs can be synthesized to address serological variants of each major neutralizing determinant for a pathogen, then combined for effective multicomponent vaccines. For example, current strategies for HIV may focus on the use of a mixture of 5-10 synthetic peptides representing variant sequences within a clade for a single principal neutralizing domain (V3 loop). An antigenic peptide library can be constructed to address not only the 5-10 major representative sequences, but also the thousands of variants within each clade that the virus is capable of producing escape from neutralizing antibodies. Second, SSALs can be synthesized to provide T-helper epitopes for a pathogen, thus overcoming the problem of immunogenetic population diversity (i.e. responders vs. non-responders) which has thus far limited the efficacy of experimental synthetic peptide vaccines. Third, SSALs can also be synthesized to include cytotoxic T cell (CTL) epitopes, again to overcome the problem of immunogenetic population diversity. Such SSALs can thereby stimulate CTL responses in broad populations, which in conjunction with a complementary serological response, is required for successful vaccines.

The "structured synthetic antigen library" differs from the random peptide libraries known in the art (see, e.g.,

Jung *et al.*, 1992, Angew. Chem. Int. Ed. Engl. 31:367). Fully random peptide libraries, produced both recombinantly and synthetically, have been used to identify receptor ligands, to determine substrate specificity, to identify epitopes. Moreover, recombinantly-produced antibody libraries, typically containing millions of randomly generated antibody Fab chains, have been described to select for high affinity antibodies.

Among the previous reports was the construction and screening of vast libraries of peptides and proteins for the ability to bind a variety of receptor molecules. This was conducted by genetic techniques fusing DNA encoding a peptide to a filamentous bacteriophage phage coat protein, to another protein for display on E. coli cells, or to a DNA binding protein such as a lac repressor. In those approaches, peptides with the desired binding properties were isolated by affinity selection (panning) on an immobilized receptor. After several rounds of panning, individual clones from an enriched clonal population were analyzed to determine the binding properties and sequences of the peptides that they contained.

Further, there are reports of synthetic peptide combinatorial libraries composed of mixtures of free peptides in quantities sufficient for use in ligand-binding assays. In one instance, the screening of a positional scanning synthetic library combined with an iterative selection and synthesis process, permitted the systematic identification of optimal peptide ligands from a "random peptide library" composed of 34 million hexa-peptides (Dooley *et al.*, 1993, Life Sci. 52:1509).

Rohren *et al.* (1993, J. Exp. Med. 177: 1713) reported a site-specific positional scanning synthetic peptide library (certain residues were fixed (anchors) and other positions had a random assortment of residues). This peptide library

was used to identify critical residues in peptide binding in the "pockets of MHC class I K^b molecules. Blake *et al.* (1992, Bioconjug Chem. 3:510) described random tetrapeptides (a set of 50,625 peptides) and hexapeptides (a set of 16,777,216 peptides) to identify residues important in the reactivity of the peptides with antibodies. The library was highly variable at each position.

Shumacher *et al.* (1992, Eur. J. Immunol. 22: 1405) reported a "limited complexity" peptide library to analyze the peptide preference of MHC class I molecules. This study taught that a singly represented T cell epitope could be detected and further suggested using the library to identify new T cell epitopes. The peptide library consisted of 8-mer or 9-mer libraries/mixtures with equimolar amounts of the variable amino acids at the multiply substituted positions and was used for epitope mapping. The peptides were based on the 9-mer Sendai virus NP peptide and the VSV-N peptide (specific for the mouse K^b locus) with the variable amino acids from correspondingly known K^b epitopes.

Hortin *et al.* (1992, Biochem. Int. 26: 731) reported a series of 5 synthetic peptide libraries of 7 amino acids in length wherein the last two carboxyl residues were totally random, i.e. included all twenty amino acids in equal proportion. One of these mixtures inhibited platelet adhesion.

Gras-Masse *et al.* (1992, Peptide Res. 5: 211) described a "mixotope" synthetic peptide library from the hypervariable gp120 V3 loop of HIV-1 for use as a vaccine. The mixotope was 22-25 amino acids in length with invariant and variant positions. The variant-position residues (also termed degenerate positions) consisted only of those amino acids which exceeded 7% abundance at a particular position and were incorporated in equimolar amounts.

The peptide library approaches described above with libraries of randomized complexity generally provide a research tool for the selection of a particular ligand or antibody to facilitate further drug design. The present invention provides "structured synthetic antigen libraries" SSAL as the key bioactive ingredient in vaccines, diagnostics and therapeutics. Such libraries provide the broad range of sequences necessitated by the strain or antigenic variation by anticipating temporal and geographic variation in antigens. The resulting SSALs are collections of peptides that maintain the character of the antigen and can simultaneously provide cross-reactivity to multiple strains of an infectious agent.

SUMMARY OF THE INVENTION

The present invention provides a library of peptides comprising a known antigenic site, epitope or ligand on a protein molecule, wherein (1) said library is optionally linked to a substantially invariant peptide domain or a core branched sequence; (2) said library consists of an ensemble domain; (3) the overall peptide length is about 8 to about 100 amino acids; (4) the sequence of said ensemble domain is represented by a consensus formula; (5) said ensemble is immunogenic, diagnostic for the epitope or is a therapeutic; and (6) the consensus formula provides that each sequence position in the ensemble contain either a single amino acid or multiple amino acids, and that when a position contains multiple amino acids, (a) the identity and ratio of those amino acids being determined by the relative prevalence of amino acids in a consensus of known variant sequences for that epitope or ligand or (b) the identity of those amino acids being determined by the amino acids present in a consensus of known variant sequences for that epitope or ligand and the ratio of amino acids is equimolar.

Furthermore the antigenic sites, epitopes or ligands of the library are from a virus, bacterium, parasite, tumor antigen, allergen or other protein antigen, from a diagnostic marker site of an infectious agent or disease, from a therapeutically valuable ligand or from a helper T cell epitope or a cytotoxic T lymphocyte (CTL) epitope. More particularly, these antigenic sites, epitopes or ligands can be from an envelope, core, NS1, NS3, NS4 or NS5 proteins of HCV; a gp120 V3 loop, gp41 envelope protein or the gp40 envelope protein of HIV; an envelope protein of HTLV I/II; an HA protein or a mutant HA domain of influenza A virus; an major outer membrane protein of Chlamydia trachomtis; a neuraminidase, SAPA or CRA site of T. cruzi; an envelope B cell site or helper T cell site of type 2 Dengue virus; an M protein of streptococcus; a carbohydrate recognition site of selectin; a promiscuous helper T cell sites; a CTL epitope from HIV gag specific for HLA-B27; or an OspA, OsaB, OspC or flagellin protein of Lyme disease. SSAL libraries of the present invention are also shown in Figs. 2-11, 13, and 15-34.

Another aspect of the invention is directed to peptide compositions containing the above libraries, including immunogenic peptide compositions.

Still another aspect of the invention provides a method of detecting antibodies associated with an infectious agent or a disease state which comprises using an effective amount of the subject peptide compositions in an immunoassay procedure. Also provided is a method of detecting infection or a disease state which comprises contacting an effective amount of the subject peptide compositions with a body fluid, tissue or tissue extract in an immunoassay procedure for a time sufficient to form a complex between said peptide composition and any antibody in said fluid, said tissue, or said tissue extract, and subjecting said complex to a

detecting means. Preferably, the immunoassay procedure is an ELISA or a PHA procedure. Similarly, the invention provides a kit for detection or diagnosis of an infectious agent or a disease state comprising a first container adapted to contain the subject peptide composition. In one embodiment, the kit is an ELISA or PHA test kit.

A still further aspect of the invention relates to a method of treating a disease which comprises administering an effective amount of the subject peptide compositions to a patient for a time sufficient to elicit an efficacious result. In this respect, the invention provides a pharmaceutical composition comprising the subject library, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

Yet a further aspect of the invention is directed to a method of vaccination which comprises administering to a mammal an amount of the subject library effective to produce a protective immune response against an infectious agent which provided the basis of the library as well as a vaccine composition therefor. Such vaccine composition contains the subject library, or an acceptable salt thereof, formulated with immunologically acceptable carrier or adjuvant.

Another aspect of the invention is directed to a method of making an open diagnostic, vaccine or therapeutic composition for a mutatable infectious agent or for a protein site of known diversity which comprises (1) aligning a collection of primary amino acid sequences for a related family of antigens, epitopes, diagnostic markers or therapeutic sites from said mutatable infectious agent or said protein site; (2) identifying invariant and variant amino acid positions in said alignment; (3) calculating a consensus formula for the sequence of a structured synthetic antigen library (SSAL); (4) preparing the SSAL from the consensus formula. The consensus formula is determined by

from examining each sequence position in the alignment to ascertain the presence of a single amino acid or multiple amino acids. When a position contains multiple amino acids, then (a) the identity and ratio of the multiple amino acids at that position in the SSAL is determined by the relative prevalence of amino acids in the alignment, (b) the identity of the multiple amino acids is determined by the amino acids present in the alignment and the ratio of the multiple amino acids is equimolar, or (c) the identity of the multiple amino acids is determined by the amino acids present in the alignment and the ratio of the multiple amino acids is weighted towards the predominant amino acid at said position. The process is repeated for each position of the sequence which can be represented by multiple amino acids.

This method for making a vaccine, diagnostic or therapeutic SSAL can be used with mutatable infectious agents from a virus, bacterium or parasite or with a protein site from a tumor antigen, allergen or other protein antigen, from a helper T cell epitope or a CTL epitope. Similarly, the family of antigens, epitopes, diagnostic markers or therapeutic sites can be selected from an envelope, core, NS1, NS3, NS4 or NS5 proteins of HCV; a gp120 V3 loop, gp41 envelope protein or the gp40 envelope protein of HIV; an envelope protein of HTLV I/II; an HA protein or a mutant HA domain of influenza A virus; an major outer membrane protein of <u>Chlamydia trachomtis</u>; a neuraminidase, SAPA or CRA site of <u>T</u>. <u>cruzi</u>; an envelope B cell site or helper T cell site of type 2 Dengue virus; an M protein of streptococcus; a carbohydrate recognition site of selectin; a promiscuous helper T cell site; a CTL epitope from HIV gag specific for HLA-B27; or an OspA, OspB, OspC or flagellin protein of Lyme disease.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates a generalized mathematical and chemical representation of a structured synthetic antigen library (SSAL).

Fig. 2 illustrates a primary amino acid sequence alignment (SEQ ID NO:1) of an antigenic NS4 peptide (47-mer) from eight representative strains of HCV. More weight (i.e. three times the listing frequency) is given to the prototype sequence used in the construction of an SSAL HCV NS4 to allow easy calculation of the ratio for each amino acid at the specified position.

Fig. 3 illustrates a primary amino acid sequence alignment (SEQ ID NO:2) of an antigenic core peptide (61-mer) from eight representative strains of HCV. An SSAL HCV core is similarly constructed based on the given antigen structure and its source sequences.

Fig. 4 illustrates a primary amino acid sequence alignment (SEQ ID NO:3) of an antigenic NS3 peptide (80-mer) from eight representative strains of HCV. An SSAL HCV NS3 is similarly constructed based on the given antigen structure and its source sequences.

Fig. 5 illustrates a primary amino acid sequence alignment (SEQ ID NO:4) of an antigenic NS5 peptide (44-mer, Sequence #1) from eight representative strains of HCV. An SSAL HCV NS5 (#1) is similarly constructed based on the given antigen structure and its source sequences.

Fig. 6 illustrates a primary amino acid sequence alignment (SEQ ID NO:5) of an antigenic NS5 peptide (41-mer, Sequence #2) from eight representative strains of HCV. An SSAL HCV NS5 (#2) is similarly constructed based on the given antigen structure and its source sequences.

Fig. 7 illustrates a primary amino acid sequence alignment (SEQ ID NO:6) of an antigenic Env peptide (55-mer, Sequence #1) from sixteen representative strains of HCV. An

SSAL HCV ENV (#1) is similarly constructed based on the given antigen structure and its source sequences.

Fig. 8 illustrates a primary amino acid sequence alignment (SEQ ID NO:7) of an antigenic Env peptide (44-mer, Sequence #2) from sixteen representative strains of HCV. An SSAL HCV ENV (#2) is constructed based on the given antigen structure and its source sequences.

Fig. 9 illustrates a primary amino acid sequence alignment (SEQ ID NO:8) of an antigenic Env peptide (40-mer, Sequence #3) from sixteen representative strains of HCV. An SSAL HCV ENV (#3) is constructed based on the given antigen structure and its source sequences.

Fig. 10 illustrates a primary amino acid sequence alignment (SEQ ID NO:9) of an antigenic Env peptide (35 mer, Sequence #4) from ten representative strains of HCV. An SSAL HCV ENV (#4) is constructed based on the given antigen structure and its source sequences.

Fig. 11 illustrates a primary amino acid sequence alignment (SEQ ID NO:10) of an antigenic NS1 peptide from eight representative strains of HCV. An SSAL HCV NS1 is constructed based on the given antigen structure and its source sequences.

Fig. 12 illustrates four types of heteromeric synthetic carriers used in the production of branched SSALs: dendritic (A), linear (B), tree-like (C), and tetrameric (D). Attached onto the core carriers is the SSAL synthesized according to a specified mathematical and chemical representation. Each dash shown within a peptide represents a particular amino acid corresponding to a conserved position in the original protein. Each 0 represents a position where peptides in the library contain any one of a fixed number of amino acids determined from the composition of the amino acid pool used in that cycle of synthesis. The

variable positions can be at any position in the peptide except branch points.

Fig. 13A illustrates a primary amino acid sequence alignment (SEQ ID NO:11, top Fig. 13B) of an antigenic gp120 Env peptide (33 mer), representing the principal neutralizing determinant (PND), the V3 domain, of the HIV-1. An SSAL HIV-1 gp120 V3 is constructed according to a previously optimized V3 framework (Fig. 13B; SEQ ID NO:11). Sixteen V3 sequences representative of geographically distinct HIV field isolates were used with more weight given to the Consensus S sequence to allow easier calculation of ratios for the various amino acids. An SSAL is constructed based on the given antigen structure and its source sequences. The SSAL HIV-1 gp120 V3 of 10^9 complexity is shown in Fig. 13C (SEQ ID NO:41). Fig. 13D depicts the neutralization titers found in the serum derived from the guinea pigs immunized with the SSALs of the indicated complexities and neutralization titers evoked by monovalent octameric V3 immunogens, $V3_{MN}$ and $V3_{91-330}$.

Fig. 14 illustrates a primary amino acid sequence alignment (SEQ ID NO:12) of a highly antigenic gp41 Env peptide (35 mer) from 50 representative strains of HIV-1. The source of the sequences is derived from Human Retroviruses and AIDS, 1993, ed. by G. Myers et al. published by Los Alamos National Laboratory, USA. An SSAL is constructed based on the given antigen structure and its source sequences.

Fig. 15 illustrates a primary amino acid sequence alignment (SEQ ID NO:13) of an antigenic gp40 Env peptide (36 mer) from 20 representative strains of Type 2 Human Immunodeficiency Virus. The source of the sequences is from Human Retroviruses and AIDS, 1993, ed. by G. Myers et al. published by Los Alamos National Laboratory, USA. An SSAL

is constructed based on the giv n antigen structure and its source sequences.

Fig. 16 illustrates a primary amino acid sequence alignment of three antigenic Env peptides from HTLV-I/II (SEQ ID NO:14, Fig. 16A; SEQ ID NO:15, Fig. 16B; SEQ ID NO:16, Fig. 16C). An SSAL is constructed based on the given antigen structure and its source sequences.

Fig. 17 illustrates a primary amino acid sequence (SEQ ID NO:17) of the critical domain from Influenza A virus hemagglutinin (HA), a surface glycoprotein capable of eliciting virus-neutralizing antibodies.

Alignment of the complete amino acid sequence of this domain, from residues 91-164, is shown for the H3N2 strains of Influenza virus A. The numbering of the amino acids begins with the first amino acid of the mature HA sequence (i.e. following cleavage of the signal sequence) as the number one. The amino acid sequences were extracted from the EMBL and the SWISSPROT Databases.

Fig. 18A illustrates a primary amino acid sequence alignment (SEQ ID NO:18) of the critical HA domain from the H1N1 strains of Influenza A virus capable of eliciting virus neutralizing antibodies.

The structural organization of the antigenic sites and the sequence sources are the same as described in Fig. 17. WIL and FPR834 strains each harbor a single amino acid deletion within this domain. The location of the deletion is marked by X. Therefore, a separate library was constructed to represent these sequences (SEQ ID NO:40; Fig. 18B).

Fig. 19 illustrates a primary amino acid sequence alignment (SEQ ID NO:19) of the critical HA domain from two H2N2 strains of Influenza A virus capable of eliciting virus neutralizing antibodies.

Fig. 20 illustrates an SSAL for a Mutant Form of the Critical Influenza Virus HA Neutralizing Domain. Synthetic peptides were made from the sequences that immediately flank the loop structure present in the HA neutralizing domain which is capable of eliciting antibody that partially neutralizes virus infectivity. Therefore, these sequences are likely to form a conformationally-dependent epitope important for eliciting a protective immune response. Antigen libraries have been constructed to mimic this antigenic structure by linking the amino acid sequences present on either side of the loop. These libraries represent deletions of the amino acid residues present between Cys residues at positions 97 and 139 in the original HA domain structure for the H3N2 type virus. The sequence of the deletion libraries was extended at the amino-terminus to include all amino acids represented between the first Cysteine residue of the loop structure and the next Cys residue on the amino-terminal side of the loop. Structured libraries representing this mutant form of the HA neutralizing domain are presented for H3N2, H1N1 and H2N2 influenza A viruses (SEQ ID NO:20, Fig. 20A; SEQ ID NO:21, Fig. 20B; SEQ ID NO:22, Fig. 20C; respectively).

Fig. 21 illustrates the amino acid sequence of the first variable domain from the major outer membrane protein (MOMP) of *Chlamydia trachomatis*. The primary amino acid sequence for the structured synthetic antigen library for VDI_a and VDI_b is shown in Figs. 21A and 21B respectively (SEQ ID NO:23 and SEQ ID NO:24, respectively). The chlamydia VDI domain sequences were extracted from the EMBL database and from Kaltenboeck, et al. (1993, J. Bacteriol. 175:487).

Fig. 22 illustrates the amino acid sequence of the fourth variable domain (VDIV) from the major outer membrane protein (MOMP) of *Chlamydia trachomatis*. The primary amino

acid sequence for the structured synthetic antigen library for $VDIV_a$ and $VDIV_b$ is shown in Figs. 22A and 22B respectively (SEQ ID NO:25 and SEQ ID NO:26, respectively). The chlamydia VDIV domain sequences were extracted from the EMBL database and from Kaltenboeck, et al. (1993).

Fig. 23 illustrates a primary amino acid sequence alignment (SEQ ID NO:27) derived from a series of 43 repeats present as a contiguous sequence of amino acid residues (596-1110) in the *T. cruzi* neuraminidase protein (NA) and a series of 14 repeats present as a contiguous sequence in the T. cruzi SAPA protein.

Fig. 24 illustrates a primary amino acid sequence alignment (SEQ ID NO:28) of a series of 21 repeats present as a contiguous sequence in *T. cruzi* cytoplasmic repetitive antigen (CRA).

Fig. 25A illustrates a primary amino acid sequence alignment (SEQ ID NO:29) of a 65 mer peptide representative of the critical neutralizing determinant of the Env protein from 33 antigenic variants of type 2 Dengue virus.

Fig. 25B illustrates a primary amino acid sequence (SEQ ID NO:30) of a 17-mer peptide representative of a helper T cell epitope present in the Env protein from 33 antigenic variants of type 2 Dengue virus.

Fig. 26 illustrates a primary amino acid sequence alignment (SEQ ID NO:31) of three type-specific sequences from the N-terminus of the streptococcal M protein. This N-terminus portion of the protein has been shown to be the site that elicits opsonic, or anti-phagocytic, antibodies.

Fig. 27 illustrates a primary amino acid sequence alignment (SEQ ID NO:32) of a T cell helper epitope from the collagen type I/II proteins.

Fig. 28 illustrates a primary amino acid sequence alignment (SEQ ID NO:33) of a carbohydrate recognition site from a group of selectin molecules.

Figs. 29A and B illustrates SSALs from two helper T cell determinants (T_h), SEQ ID NO:34 and SEQ ID NO:35, respectively. The general feature of such SSALs is that the degenerate T_h sequence is separated from the LHRH sequence, EHWSYGLRPG, by two Gly residues. The Gly residues act as a spacer between the T_h epitope and the LHRH sequence.

Fig. 30 illustrates a primary amino acid sequence alignment (SEQ ID NO:36) of an HLA-B27 restricted self peptide antigen. The antigen sequences are derived from histone H3, HSP89α, HSP89β, HEF 2, Helicase, ribosomal protein and L28 (ribosomal protein).

Fig. 31 illustrates a primary amino acid sequence alignment (SEQ ID NO:37) of a HLA B-27 restricted HIV GAG peptide CTL antigen. The GAG CTL antigen sequences represent those derived from HIV-1 MN, HIV-1 ELI, HIV-2 ROD and multiple HIV patient isolates.

Fig. 32 illustrates two $SSAL_{ospA}$ antigens by chemical and mathematical representations #1 and #2 (SEQ ID NO:38 and SEQ ID NO:39, respectively). The libraries were designed according to the primary sequence alignment provided for 12 isolates of European and North American origin by Wallich et al; 1992 (ref. cited in Example 16). These included five isolates of genospecies *B. burgdorferi* sensu stricto (ospA serotype 1), five isolates of *B. garinii* (one of ospA serotype 4, three of ospA serotype 6, and one of an unexpressed ospA), and sequences from two isolates of the V_{461} genospecies (ospA serotype 2). The amino acid ratios for variable positions were adjusted to give the *B. burgdorferi* sensu stricto sequences a weight of 50%, so as to account for the greater than 90% prevalence of this genospecies in North America. *B. garinii* and V_{461} isolates were weighted to approximately 25%, to account for the distributions of these genospecies in Europe.

$SSAL_{ospA}$ #1 includes ospA amino acids 175-234 and #2 is an overlapping sequence of residues 210-273. The overlap in the two libraries is to ensure that each contains B and T cell antigenic determinants, both being required for protective immunity.

Fig. 33 illustrates an SSAL OspC antigen by chemical and mathematical representations (SEQ ID NO:42).

Fig. 34 illustrates an SSAL flagellin antigen by chemical and mathematical representations (SEQ ID NO:43).

DETAILED DESCRIPTION OF THE INVENTION

In this invention, a "structured synthetic antigen library" or SSAL corresponds to the antigenic region of a protein from an infectious agent or from another biologically important system, a diagnostic marker or a variable therapeutic domain of a protein. The design of an SSAL covers the broad range of sequences embraced by geographical or population differences and temporal variation in these molecules.

A "structured synthetic antigen library" or SSAL is composed of an ordered set of from at least 3 to several million different but related peptides having sequences imposed upon an invariant structural framework capable of maintaining the antigenicity, diagnostic value or therapeutic bioactivity of that site. Fig. 1 provides a mathematical and chemical formula for the SSAL, which formulation describes the relative ratio of amino acids at each position in the SSAL. AA_1 to AA_i represent the amino acid sequence from N- to C-terminus of the library, j varies from 1 to n where n represents the number of possible different amino acids known at the ith amino acid position. Following this formula, a SSAL can be prepared in a single synthesis.

The sequence of the SSAL is determined by aligning the primary amino acid sequences of a related family of antigens, markers or diagnostics and identifying the invariant and variant loci within the alignment. The invariant loci generally represent the structural framework of the SSAL. The degeneracy within the SSAL is determined by the loci within the alignment that harbor different amino acid residues relative to an arbitrary prototype sequence. After determining which amino acids are to be at each position, the degree of degeneracy for the multiresidue containing positions in the SSAL library is determined from the number of variants each individual amino acid represents. The SSAL is then synthesized with single amino acids at the invariant positions and with the requisite degeneracies at variant positions. Thus in a simple manner, the specific amino acids and their frequency of appearance at each position within the SSAL is defined by the primary sequences of the different antigens or molecules that make up the alignment.

In another embodiment, amino acids not known to exist at a variant locus within aligned epitopes can also be incorporated into the library. For example, all residues of a specific class of amino acids (e.g. hydrophobic, charged, neutral or polar) can be incorporated at a variable locus. The class of amino acids incorporated at a specific locus can be determined by the predominant type of amino acid found at that variant locus within the antigen alignment.

In broad application, the sequence of any SSAL can be defined by homologous portions of existing primary amino acid sequences from a family of proteins or peptides of known biological significance. The sequences encoding the family of proteins or peptides encoding the target antigenic site, diagnostic marker or variable therapeutic domain of a protein or peptide may represent: 1) alternate alleles of

the same gene; 2) genes that are evolutionarily related as deduced from sequence homology, but isolated from different strains or species of similar organisms; 3) genes encoding structurally similar proteins with similar function but isolated from very different organisms, or; 4) repeated sequences within the same gene, or repeats from different genes within the same organism or repeats within homologous genes from different, but related organisms. The order within any SSAL is provided by invariant amino acid residues that are defined as the framework sequence of the selected protein or peptide.

The SSAL can be from about 8 to about 100 amino acids in length depending on the particular site and preferably from about 10 to about 100. The overall length of the SSAL can be minimized to provide the fewest residues necessary to elicit the desired biological response. Similarly, extra residues can be added to the ends of the SSAL. For example, KKK can be added at the amino terminus to increase peptide solubility, cysteine can be added to facilitate directed coupling to carrier molecules, and methionine can be added for cyanogen bromide cleavage if necessary.

Moreover, the SSAL can be a domain within a peptide or can have other antigenic, diagnostic or therapeutic sites attached to it. The SSAL can be attached to a core sequence for facile delivery. These core sequences include dendritically branched cores, linear array type branched cores or randomly branched cores (e.g. poly-L-lysine). The branched cores can be composed of an amino acid or an amino acid analog having two amino groups and one carboxyl group, each group capable of forming a peptide bond linkage. Preferably such amino acids are lysine or a lysine analog such as ornithine. The amino acid analog can be an α-amino acid, a ß-amino acid, or any other either natural or non-natural amino acid with two amino groups and one carboxyl

group available for forming peptide bonds. Preferred branched peptides of the invention are dimers, tetramers and octamers, especially those having a branching core structure composed of lysine. Similarly, the branched cores can contain other residues interspersed among the branching residues as depicted, for example, in Fig. 12.

If the variant sequences comprise insertions and/or deletions of amino acids relative to the prototype sequence, separate SSALs may need to be prepared to fully accommodate the various strains. Such SSALs preserve the immunoreactivity or bioactivity of the desired antigen, marker or therapeutic.

The subject SSALs can also be used to form conjugates, i.e., the SSAL, either in branched or linear form can be coupled directly or indirectly, by methods known in the art, to carrier proteins such as bovine serum albumin (BSA), human serum albumin (HSA), or to red blood cells or latex particles.

As used herein, natural amino acids are the 20 amino acids commonly found in proteins (i.e. alanine, aspartic acid, asparagine, arginine, cysteine, glycine, glutamine, glutamic acid, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, tryptophan and valine). The natural amino acids also include the D- and L- forms of such amino acids.

As used herein "unnatural amino acids" include both D- and L- forms of any other amino acids whether found in a protein, whether found in nature or whether synthetically produced. Unnatural amino acids can include, but are not limited to, ß-alanine, ornithine, norleucine, norvaline, hydroxyproline, thyroxine, gamma-amino butyric acid, homoserine, citrulline and the like.

The SSALs of the present invention are thus from antigenic sites of viruses, bacteria, parasites, tumor

antigens, allergens or other protein antigens as well as the relevant diagnostic marker sites of these infectious agents or diseases. Likewise the SSALs include helper T cell epitopes, CTL epitopes or other variable immune epitopes. By way of example, SSALs can be prepared from the envelope, core, NS1, NS3, NS4 or NS5 proteins of HCV; the gp120 V3 loop, gp41 envelope protein or the gp40 envelope protein of HIV; the envelope protein of HTLV I/II; the HA protein and a mutant HA domain of influenza A virus (including the H3N2, H2N2 and H1N1 strains); the major outer membrane protein of _Chlamydia_ _trachomitis_ (including the VDI and VDIV sites); the neuraminidase (NA), SAPA and CRA sites of _T_. _cruzi_; the envelope B cell site and helper T cell site of type 2 Dengue virus (or other types); the M protein of streptococcus; the carbohydrate recognition site of selectin; promiscuous helper T cell sites; the CTL epitope from HIV gag specific for HLA B-27; and the OspA, OspB, OspC and flagellin proteins of Lyme disease (from _Borrelia_ _burgdorferi_ and _B_. _garinii_. The SSALs of this invention further include but, are not limited to, those specifically provided in the Examples, s and Sequence Listings.

The SSALs are prepared by chemical synthesis using standard techniques well known in the art such as the solid-phase synthetic route pioneered by Merrifield. The coupling of multiple amino acids at a given position is accomplished by providing a mixture of the desired amino acids at the appropriate ratios. If necessary the ratio of amino acids in the mixture can be varied to account for different coupling efficiency of those amino acids. The SSALs can also be produced by standard recombinant DNA technology, pooled in batch, isolated if necessary and used as provided in accordance with the invention. Preferably, SSALs are prepared by chemical synthesis.

The peptide compositions of the present invention can be composed of one or more SSAL. Preferably such compositions contain from one to about 30 SSALs depending on the application.

SSALs in accordance with this invention are synthesized and tested to determine immunoreactivity, elicitation of neutralizing antibodies, diagnostic utility, therapeutic value and the like of the SSAL as described in the Examples, by ELISA, by PHA, by bioassay or any other technique appropriate to the source of the SSAL. For example, the efficacy of the HCV SSALs in detecting and diagnosing NANBH and HCV infection is determined by testing it for reactivity with serum specimens with known immunoreactivity for HCV. Such serum panels are commercially available. The strategy for serological validation, naturally, depends on the expected characteristics of the target epitopes. For example, universal immunodominant epitopes, such as the gp41 transmembrane peptide of HIV-1, can be screened by a single representative serum sample from a patient known to be infected with the virus. Epitopes which are not recognized by all infected individuals, or those for which antibody is produced late or only transiently, and especially epitopes which give rise to neutralizing antibodies, may need to be screened by large panels of sera. Both methods of screening can be employed in the present invention to assess the SSALs for selectivity and sensitivity.

Based on the immunoreactivities of the SSALs, they are useful in a vaccine composition to treat or prevent the infection caused by the infectious agent from which they are derived. These vaccine compositions containing one or more SSALs, alone or when coupled to a carrier or polymerized to homo- or hetero-dimers or higher oligomers by cysteine oxidation, by induced disulfide cross-linking, or by use of homo- or hetero-functional multivalent cross-linking

reagents, can be introduced into normal subjects to stimulate production of antibodies. Similarly the subject SSALs can be formulated in a vaccine composition using adjuvants, pharmaceutically-acceptable carriers or other ingredients routinely provided in vaccine compositions. Such formulations are readily determined by one of ordinary skill in the art and include formulations for immediate release and for sustained release, e.g., microencapsulation. The present vaccines can be administered by any convenient route including subcutaneous, oral, intramuscular, intravenous, or other parenteral or enteral route. Similarly the vaccines can be administered as a single dose or divided into multiple doses for administration.

The vaccine compositions of the instant invention contain an immunoeffective amount of the SSAL to treat or prevent the target infection. Such compositions in dosage unit form can contain about 0.1 μg to about 1 mg of the peptide (or mixture of peptides) per kg body weight. When delivered in multiple doses, the dosage unit form is conveniently divided into the appropriate amounts per dosage.

The SSAL peptide compositions prepared in accordance with the present invention can be used to detect or diagnose a target infection by using them as the test reagent in an enzyme-linked immunoadsorbent assay (ELISA), an enzyme immunodot assay, a passive hemagglutination assay (e.g., PHA test) or other well-known immunoassays. In accordance with the present invention, any suitable immunoassay can be used with the SSALs. Such techniques are well known to the ordinarily skilled artisan and have been described in many standard immunology manuals and texts, see for example, by Harlow et al. (1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 726 pp. In a preferred embodiment, the immunoassay is an ELISA

using a solid phase coated with the peptide compositions of the present invention. ELISA techniques are well known in the art. In another preferred embodiment the immunoassay is a PHA assay.

The immunoassays of the present invention are used to screen body fluids and tissues for the presence of the target infectious agent or diagnostic marker. The body fluids which can be subjected to such screening include blood and blood fractions (e.g. serum), saliva, or any other fluid which contains antibodies specific for the target infectious agent or pathogen.

The examples serve to illustrate the present invention and are not to be used to limit the scope of the invention.

Example 1

EFFECTIVE DETECTION OF HCV ANTIBODIES BY NS-4 SSAL IN PATIENTS FROM GEOGRAPHICALLY DISTINCT REGIONS

An HCV NS4 antigen library was designed according to the sequence listing shown in Fig. 2. The HCV NS4 SSAL was synthesized as linear peptides using the solid phase peptide synthesis strategy employing Fmoc chemistry with standard side chain protecting groups. Protected amino acids were added sequentially during the synthesis process from C- to N-terminus according to the mathematical and chemical representation of the library as shown in Fig. 2B. The protected amino acid reagent added at each of these cycles consists of a collection of amino acid types, where the total concentration of amino acids is identical to that of non-variable positions, but the ratio of each type was set by an algorithm specific to the application of the individual SSAL. The deprotection and cleavage procedures were carried out by standard procedures applicable to the Fmoc chemistry of the synthesis. As controls, two NS4 peptides with corresponding HCV-1 and HCV J-8 sequences as listed in Fig. 2 were also synthesized as linear peptides for comparison of their relative HCV immunoreactivity.

Wells of a microtiter plate were coated with NS-4 HCV-1 peptide, the J-8 variant analogue, or the NS-4 SSAL at 5 µg/mL. The NS-4 SSAL peptide was selectively substituted in 27/47 (57%) of its amino acids. In five of these substituted positions, the substituted amino acid corresponding to the J-8 sequence represents only 10% of the total. Specimens from two groups were tested: Japanese hemodialysis patients, a group at high risk for hepatitis C (HCV) infection, for whom recombinant C-100 assay data was available, and plasma donors confirmed seropositive for HCV by recombinant immunoblot (RIBA-II). Two samples that were non-reactive on the HCV-1 peptide reacted with both the J-8

peptide and the SSAL peptide (Table 1a), despite the minority representation of the J-8 sequence in the SSAL peptide. Both samples were non-reactive with the HCV-1 recombinant C-100 protein, indicating that additional antigenic regions that may exist in the C-100 protein do not contribute toward recognition of the Japanese patient antibody. The SSAL NS-4 peptide also detected samples that reacted with the HCV-1 peptide and recombinant C-100 protein but not with the J-8 peptide (Table 1b). In five samples in which both HCV-1 and J-8 peptides were reactive (Table 1c), the response to the SSAL peptide was greater than to either of the single strain peptides derived from HCV-1 or J-8. In the case of Japanese dialysis patient No. 97, the absorbance on the SSAL peptide was twice as high as on either HCV-1 or J-8 peptide.

The results obtained from the above evaluation of critical serological samples with HCV immunoassays demonstrated a clear advantage in using the SSAL antigen design approach for the detection of antibodies to infectious agents, particularly those with high variability such as HCV.

Table 1
Detection of HCV Antibodies by NS-4 SSAL in HCV-infected Japanese Dialysis Patients and US Plasma Donors

Specimen Source	Sample ID	rC100 EIA	RIBA-II[1]				NS-4 Peptide (A_{492})		
			511	c100	c33	c22	HCV-1	J-8	SSAL
a. Japan	97	Neg	ND	ND	ND	ND	0.022	1.155	0.635
US	9	Neg	0	0	4	4	0.051	0.313	0.233
b. US	1	Pos	3	4	4	4	1.264	0.043	1.206
US	3	Pos	2	2	4	4	1.073	0.066	1.026
US	21	Pos	4	4	4	0	2.273	0.089	2.398
c. Japan	57	Pos	ND	ND	ND	ND	0.520	0.521	1.096
US	5	Pos	4	3	4	4	2.009	1.084	2.507
US	16	Pos	4	4	4	2	1.583	1.301	2.120

[1]In a recombinant immunoblot assay (RIBA), antibody reactivity to each of the HCV recombinant proteins 511, c100, c33 and c22 is scored from 0 to 4 with 4 being the strongest.

ND: not done

2008Z-WO
10/25/94

Example 2

SSAL HCV PEPTIDES CAN BE EFFECTIVELY FORMULATED INTO AN HCV IMMUNOASSAY WITH IMPROVED SENSITIVITY AND SPECIFICITY OVER ASSAYS USING INDIVIDUAL HCV-1 PEPTIDES

Five HCV antigen [NS4, core, NS3, NS5(#1) and NS5(#2)] libraries were designed according to the sequence listing shown in Figs. 2, 3, 4, 5 and 6 respectively. All five HCV antigen SSALs were synthesized as linear peptides using the solid phase peptide synthesis strategy employing Fmoc chemistry with standard side chain protecting groups according to the mathematical and chemical representation of each of the libraries as illustrated in Figs. 2, 3, 4, 5 and 6. All deprotection and cleavage methods were identical to standard procedures applicable to the Fmoc chemistry of the synthesis.

SSAL peptides NS4, core, NS5(#1), NS5(#2), and NS3 were coated onto wells of a microtiter plate at a concentration of 2, 1, 0.5, 0.5 and 5 μg/Ml, respectively. The NS3 SSAL peptide was employed as a conjugate to bovine serum albumin (BSA). The reactivity of the mixed SSAL peptides was compared to the standard HCV peptides under the same formulation. Sensitivity was evaluated using commercially available low titer and mixed titer HCV panels from Boston Biomedica, Inc. In five samples from panels 102 and 202, the SSAL mixture of peptides provided greater sensitivity than did the mixture of individual HCV-1 peptides (Table 2). According to RIBA-II analysis, the reactivity of the five samples that had higher absorbance on the SSAL mixture of peptides varied: patterns of core only, NS-3 only and mixtures of NS-3/NS-4 or NS-3/core were all represented. Specificity of the two formats tested over 264 random plasma samples was equivalent. One well- characterized false positive sample, detected by the HCV-1 mixture but not by

RIBA-II, was negative on the SSAL peptide mixture. The results demonstrated improved HCV immunoassay performance using SSAL-based HCV peptides.

Example 3

BRANCHED SSALS REPRESENTING THE HIGHLY VARIABLE YET ANTIGENIC REGIONS OF THE HCV ENV AND NS1 PROTEINS ARE USED AS THE KEY INGREDIENTS IN A POLYVALENT VACCINE FOR THE PREVENTION OF INFECTION BY HIGHLY DIVERGENT GLOBAL STRAINS OF HCV

HCV has been associated with liver disease, both acute and chronic, and with the development of liver cancer. Chronic liver disease occurs in at least 50% of infected individuals, and up to 20% of these go on to develop cirrhosis. Variation in immune responsiveness to HCV may account for the variation in outcome associated with HCV infection.

In this HCV vaccine development approach, branched SSALs derived from highly variable yet antigenic regions of the HCV Env and NS1 proteins are used in a polyvalent vaccine. More specifically, five HCV SSALs derived from both the Env and NS1 protein regions are designed according to the sequence listing shown in Figs. 7, 8, 9, 10 and 11. Analogues of the four Env peptides and one NS1 peptide with sequences corresponding to the HCV J-1 strain were previously found to be antigenic, in that they reacted with sera from HCV- infected individuals. In this HCV vaccine design, all five HCV Env and NS1 SSALs are synthesized as branched peptides onto a heteromeric dendritic branched carrier as shown in Fig. 12a, using the solid phase peptide synthesis strategy employing Fmoc chemistry with

Table 2
Improved HCV Assay Sensitivity with Low Titer HCV Antibodies by Employing

HCV SSAL as the Solid Phase Antigens

Sample Number	HCV Peptide Antigen[1]		HCV SSAL Antigen[2]		Abbott 2.0 HCV Ratio	Ortho 2.0 HCV Ratio	Ortho RIBA 2.0				
	Abs	Ratio	Abs	Ratio			5-1-5	c100-3	C33C	C22-3	Result
PHV-102-05	0.220	1.41	0.836	4.26	1.1	2.7	-	-	-	1+	IND
PHV-202-09	0.587	3.76	1.413	7.21	4.7	4.4	3+	1+	4+	-	POS
PHV-202-10	0.260	1.67	0.396	2.02	2.7	2.8	-	-	3+	-	IND
PHV-202-14	1.172	7.51	2.285	11.66	4.7	4.4	-	-	3+	4+	POS
PHV-202-22	0.853	5.46	1.543	7.87	4.7	4.4	-	-	4+	-	IND

[1]Five HCV peptides (1 from core, 1 from NS4, 2 from NS5 and 1 from NS3 regions) were used to formulate an HCV enzyme immunoassay.
[2]Five HCV antigen libraries corresponding to each of the five antigenic peptides were used to formulate an SSAL HCV enzyme immunoassay.

standard side chain protecting groups. Protected amino acids are added sequentially during the synthesis process from C-toN-terminus according to the mathematical and chemical representation of each of the libraries. All deprotection and cleavage methods are carried by standard procedures applicable to the Fmoc chemistry of the synthesis.

After the synthesis of five HCV Env/NS1 SSALs as branched peptides, they are formulated with appropriate vaccine vehicles and adjuvants and tested for their capacity to induce HCV neutralizing antibodies in laboratory animals which will block the growth of HCV isolates in vitro. Following demonstration of neutralizing antibody production in vivo, the SSALs can be formulated into microparticles or other appropriate vaccine delivery systems for efficacy testing in human volunteers at risk for HCV infection.

Example 4

SYNTHETIC AIDS VACCINE FOR THE PREVENTION OF HIV-1 INFECTION USING SSAL REPRESENTATIVE OF UNIQUE AMINO ACID SEQUENCE FRAME DERIVED FROM THE GP120 V3 REGION

The third hypervariable region (V3) of the human immunodeficiency virus-1 (HIV-1) external envelope glycoprotein (gp120) contains a cysteine loop, the V3 loop, that constitutes the principal neutralizing determinant (PND). To enhance the immunogenicity of PND peptides, the V3-derived peptides have been presented as radial eight-branched octamers based on an optimized amino acid sequence frame. A heptalysyl core bearing eight reactive NH_2-termini served as a carrier onto which eight identical V3 peptides were attached by solid-phase synthesis. Significant titers of V3 reactive neutralizing antibodies were observed after immunization with this branched peptide vaccine (Wang et al. 1991, Science 254: 285). This V3 peptide immunogen thus serves as a promising approach for an AIDS vaccine. However, two major problems

were encountered subsequently with immunogenicity during the development of such a vaccine. The sources of the problems, the hypervariability of HIV-1 V3 and the genetic restriction of peptide vaccines by the human hosts, served to render a simple peptide insufficiently immunogenic for application in a peptide-based vaccine.

First, the hypervariability of the HIV-1 gp120 V3 PND presented a major challenge. Although immunizations with a mixture of the branched octameric V3 peptides carefully selected from the HIV-1 V3 sequences shown in Figure 13A resulted in broadly crossreactive anti-V3 neutralizing antibodies capable of reacting with and neutralizing isolates from each major clade of HIV-1 (e.g. Clades A, B, C, D, E), intra-clade variation, e.g., variation within clade A (e.g. A-1. A-2..., A-n) still presented a significant obstacle to an effective vaccine.

The second major problem, the genetic restriction of the cellular immune response to a peptide immunogen, became evident during a phase I clinical trial using the monovalent branched octameric $V3_{MN}$ peptide in a prototype vaccine. Only a subpopulation of the vaccinees responded to the V3 immunogen after a six-month three-dose protocol. This narrow range in the responsiveness of humans to the peptide-based vaccine was a manifestation of the simplicity of the immunogen design: The monovalent peptide, comprised of eight identical $V3_{MN}$ peptides attached to heptalysine core, had an insufficient number of T cell epitopes for broad responsiveness in the genetically diverse human population.

More diversity associated with the V3 sequences is needed to allow the synthetic immunogen to simultaneously induce neutralizing antibodies to the globally diverse HIV-1 V3 spectrum (e.g. inter-clade and intra-clade variation) and to elicit responses in the majority of genetically variable human subjects.

Because of the large numbers of different viruses expressing different V3 loop primary sequences, the synthesis of individual sequences, either by synthetic or by recombinant DNA means becomes impractical, if not impossible. Clearly, the production of an array of V3 immunogens as an SSAL, by a single peptide synthesis, provides a useful solution to this problem. Furthermore, a V3 SSAL can impart an additional advantage to a peptide-based vaccine. By the insertion of amino acid mixtures at the variable loci within the V3 epitope array, a V3 SSAL immunogen can elicit protective immune responses specific to HIV-1 types that have not yet been produced by the readily mutable virus, thereby anticipating viral evolution and preventing evasion of the immune response through escape mutation.

A branched SSAL with sequences derived from the previously characterized V3 region of the HIV-1 gp120 protein can be synthesized onto a standard heptalysyl core or any one of the four types of heteromeric branched core resins shown in Fig. 12 using the solid phase Fmoc chemistry with standard side chain protecting groups. For example, a V3 SSAL was synthesized onto a heptalysyl core resin according to the mathematical and chemical representation of the library as shown in Figs. 13A and B. The protected amino acid reagent added at each of these cycles (representing a variant position), consists of a collection of amino acid types, where the total concentration of amino acids at the variant position is identical to that of non-variable positions, but the ratio of each type is set by an algorithm specific to the application of the individual SSAL. All deprotection and cleavage methods were carried out identically to standard procedures for the Fmoc chemistry of the synthesis.

The SSAL HIV-1 V3 immunogen was evaluated by immunizations followed by assays for serological and cellular immune activities; and, in demonstration of the value and

flexibility for the complexity of the V3 SSAL, complexity was varied and its influence on immunogenicity was observed.

<u>a) Serology of response to the SSAL HIV-gp120 V3 immunogen.</u> Following the synthesis of the SSAL HIV-gp120 V3, the peptide library was formulated in 0.5% alum and injected into four guinea pigs at 100 μg per dose. Three and six weeks after the first injection, sera from the animals were tested by a panel of anti-peptide ELISAs (enzyme-linked immunosorbent assays) for their immune reactivities against the immunizing V3 peptide SSAL as well as for reactivities to other individual V3 peptides whose sequences (Figure 13A) contributed to the construction of the V3 SSAL. As shown in Table 3, the SSAL HIV-gp 120 V3 immunogen induced significant antibody responses (Log_{10} titer > 4.0) to itself, the "V3 library", and more significantly, demonstrated broad reactivity to monovalent V3 peptides with sequences derived from geographically distinct worldwide isolates. In this experiment, the branched SSAL HIV-1 gp120 V3 demonstrated potent immunogenicity and elicited highly crossreactive V3 antibodies.

Another aspect of the serological reactivity evoked by the V3 SSAL immunogen is neutralization activity, i.e., the capability of the anti-sera to neutralize or inactivate infectious virus. The ability to elicit viral neutralization by antibodies is an important component of a protective vaccine against HIV infection. Neutralization activity was determined by the Rapid Microplaque Assay (MT-2 assay) decribed by Hanson et al. (1990, J. Clin. Microbiol. 28:2030) in which the infectivity of HIV-1 (preincubated with neutralizing serum) is determined on HIV-sensitive MT-2 cells. The particular MT-2 assay was a stringent assay in which the endpoint was the antibody dilution at which 90% of input virus, North American HIV-1 isolate MN of Clade B, is neutralized. This endpoint is expressed as the "$MT\text{-}2MN_{90}$".

The pooled antiserum from the V3 SSAL-immunized guinea pigs had an MT-2MN$_{90}$ of 1:2430. This titer was comparable to the MT-2mn$_{90}$ of ≥1:2430 achieved by guinea pigs administered comparable doses of an analogous monovalent peptide immunogen representing the V3 sequence for HIV-1 MN only. These titers were comparable dispite the minute proportion of V3$_{MN}$ sequence present in the SSAL immunogen (10^{-11}%) compared to the proportion of V3$_{MN}$ in the monovalent formulation (100%).

Neutralization activity for the anti-SSAL V3 antibodies in the pooled guinea pig anti-sera was also determined by a stringent assay, the infectivity reduction assay (IRA), described in White-Scharf et al. (1993, Virology 192:197), which measures neutralization activity against primary field isolates. IRA neutralization activity, expressed as infectious units (IU) blocked 100% by 1:10 dilution of serum, was 10 IU against the Zambian HIV-1 isolate Zam200370, a field isolate of Clade C, which was not contained in the V3 SSAL. This cross-neutralization activity illustrates the library's capacity for broad immunogenicity.

<u>b) CTL response to the SSAL HIV-gp120 V3 immunogen.</u> Cell-mediated immunity, including cytotoxic T lymphocyte response (CTL) is an important component of the protective immune response to HIV-1. CTL response results in the destruction of HIV-infected cells. The CTL response to a V3 SSAL immunogen is useful both in a prophylactic vaccine by preventing the transmission of infected cells, and in a therapeutic vaccine by reducing the viral load in HIV-infected individuals. CTL responsiveness to a peptide immunogen in humans is determined by how well the immunogen is recognized by the Class I HLA antigens of the genetically diverse human population. Antigens of greater complexity are more likely to be recognized by the Class I antigens of a greater part of the population. The capability of the V3 SSAL to be

recognized by a Class I major histocompatibility complex and evoke CTL was demonstrated in a mouse model system.

The immunizations were accomplished by the following protocol: The V3 SSAL branched peptide was formulated into polylactide-co-glycolide microparticles that are suspended in phosphate-buffered saline (O'Hagan et al. 1991, Vaccine, 9:768) and injected intraperitoneally into a group of female Balb/c mice (6-8 weeks old), three mice per group, at 100 μg of peptide in 0.5 ml of microparticulate vaccine per dose. Booster immunizations were administered on days 12 and 21. Animals were sacrificed on day 27 and splenocytes collected and restimulated with a monovalent linear peptide of sequence I-G-P-G-R-A-F-Y-T-T, corresponding to a portion of the V3 domain of HIV-1 MN that was present in the SSAL as an individual sequence. The same peptide was also used to sensitize target cells. The cultured splenocytes were restimulated on days 7 and 14 after sacrifice and assayed for CTL 6 days later by ^{51}Cr-release assay (Hioe et al. 1990, J. Virol. 64:6246). Target cells were ^{51}Cr-labeled A20.1-11 (H-2^d) cells pulsed with the above monovalent linear peptide or irrelevant peptide and no peptide as controls. A highly significant level of CTL immune reactivity specific for the $V3_{MN}$ peptide was observed, as represented by per cent ^{51}Cr released of (% specific lysis) 51%, 42%, and 21% for respective Effector:Target ratios of 150, 50, and 17. By comparison, CTL results with unsensitized target cells and for target cells coated with an irrelevant peptide were less than 3%. Splenocytes cultured from placebo-immunized mice displayed only background levels of ^{51}Cr release.

<u>c) Optimization of SSAL complexity.</u> The V3 SSAL evaluated above in parts a and b, with a complexity of 10^{13}, was evaluated for immunogenicity relative to that of a series of V3 libraries of diminishing complexities. Complexity is a

highly significant parameter for immunogenicity as it determines not only the number of different peptides that is delivered by an SSAL vaccine, but also the dose of any individual peptide in the SSAL.

Four V3 SSAL immunogens of increasingly simpler compositions were synthesized, of complexities 10^{12}, 10^9, 10^6, and 10^3. These were used to immunize groups of five guinea pigs, as described in part a of this Example. The composition of the V3 SSAL of 10^9 complexity is illustrated in Figure 13c. Neutralization activities (against HIV-1 MN) elicited by those libraries were determined for the pooled anti-sera of each group at week 6 as MT-2MN_{90} values and results were compared to the neutralization titer of the anti-sera to the original V3 SSAL (Figure 13B) and to neutralization titers evoked by monovalent branched V3 peptide immunogens. Neutralization activities are plotted in Figure 13D and show a strongly significant neutralization titer for the SSAL of 10^9 complexity (MT-2MN_{90} 1:9081) and the other SSALs elicited titers comparable to the results obtained from immunization with monovalent $V3_{MN}$ immunogen. Immunization with a monovalent V3 immunogen other than MN that was also present in the libraries evoked poorer cross-neutralizing activity against HIV-1 MN than was evoked by any of the SSAL immunogens regardless of library complexity.

The V3 SSAL immunogens of complexities 10^{13} and 10^3 were compared for immunogenicity by CTL assays against $V3_{MN}$-pulsed cells. Mice were immunized with these two libraries and evaluated for CTL by the protocols described in part b. At Effector:Target cell ratio of 150, results were 53% specific release for the CTL elicited by SSAL of 10^{13} sequences and 79% specific release for the SSAL of 10^3 sequences, strongly suggestive of the influence of complexity on the

Table 3

Branched SSAL HIV gp 120 V3 Demonstrated Potent Immunogenicity and Elicited Highly Crossreactive V3 Antibodies Required of a Synthetic AIDS Vaccine

Clades	V3 Library	B	S	D	B	B	E	B		A	C		B	B	D
V3 Sequence Source	UBI	BAL-1	Lab.	Thai, Uganda	HN	SC	Thai	Thai	IvoCot	Dijibod	Indi, Zar	Rwanda	Haiti	WMJ2	Uganda
Code of Octamer	620	369M	468	471	200H	281M	481	482	485	469	470	473	478	282M	486
Cross Reactivity Log_{10} Titer (3 wks)	4.015	3.500	4.165	3.529	3.670	3.461	3.512	3.995	3.490	3.630	3.130	3.039	2.568	3.675	2.990
Cross Reactivity Log_{10} Titer (6 wks)	4.749[a]	4.200	4.628	4.077	4.346	4.229	4.339	4.514	4.204	4.087	4.012	3.795	3.653	4.467	4.088

Sera obtained from 6 wks G.P. bleeds also reacted strongly with V3-library based branched peptides presented on a tetrameric linear array as shown in Fig. 12(d).

immunogenicity of HIV-1 gp120 V3 SSAL immunogens, and a clear illustration of the potent immunogenicity of a V3 SSAL.

Example 5

DETECTION OF HIV-1 ANTIBODIES WITH AN SSAL DERIVED FROM A HIGHLY ANTIGENIC REGION OF THE GP41 TRANSMEMBRANE ENVELOPE PROTEIN OF HIV-1

The gp41 transmembrane envelope protein of the human immunodeficiency virus, Type 1 (HIV-1) contains an immunodominant antigenic region useful for detection of antibodies to HIV-1. Although this antigenic region is highly conserved, variant amino acid substitutions have been observed in field isolates.

An HIV-1 gp41 antigen (35 mer) library was designed according to the sequence listing of 50 field isolates shown in Fig. 14. The HIV-1 gp41 SSAL was synthesized as linear peptides using solid phase Fmoc chemistry. Wells of a microtiter plate were coated with the SSAL HIV-1 gp41 at 1 µg/mL. Confirmed HIV-1 positive specimens from geographical diverse areas were tested with the HIV-1 gp41 SSAL and found to have strong seroreactivities.

Due to the inherent design advantages associated with the SSAL, such a gp41 antigen preparation is useful for comprehensive HIV-1 antibody detection.

Example 6

DETECTION OF HIV-2 ANTIBODIES WITH AN SSAL DERIVED FROM A HIGHLY ANTIGENIC REGION OF THE GP40 TRANSMEMBRANE ENVELOPE PROTEIN OF HIV-2

The gp40 transmembrane envelope protein of the human immunodeficiency virus, Type 2 (HIV-2) contains an immunodominant antigenic region useful for detection of antibodies to HIV-2. Although this antigenic region is

highly conserved, variant amino acid substitutions have been observed in field isolates.

An HIV-2 gp40 antigen (36mer) library is designed according to the sequence listing of 20 field isolates shown in Fig. 15. The HIV-2 gp40 SSAL is synthesized as linear peptides using the solid phase Fmoc chemistry. Wells of a microtiter plate can be coated with the SSAL HIV-2 gp40 at 1μg/mL. Confirmed HIV-2 positive specimens, mostly from West Africa are tested against these SSALs.

Due to the inherent design advantages associated with the SSAL, such an HIV-2 gp40 SSAL antigen preparation is thus useful for comprehensive HIV-2 antibody detection in the event of antigenic drift.

Example 7

SIMULTANEOUS DETECTION OF HTLV I/II ANTIBODIES WITH SSALS DERIVED FROM THE HIGHLY ANTIGENIC REGIONS OF THE EXTERNAL PORTION (GP46) AND TRANSMEMBRANE PORTION (GP21) OF THE HTLV ENVELOPE PROTEIN

HTLV-I and HTLV-II are closely related viruses with extensive serologic cross-reactivity. HTLV-I is the causative agent of adult T-cell leukemia and a degenerative neurological disorder, HTLV-I-associated myelopathy (HAM) or tropical spastic paraparesis. Disease association with HTLV-II is less well-established, but HTLV-II virus has been identified in patients with a neurological disease similar to HAM.

HTLV-I and HTLV-II can be transmitted by blood transfusion. Immunodominant epitopes have been identified in the gp46 external portion and gp21 transmembrane portion of the HTLV-I envelope that are useful for detection of antibodies to HTLV-I.

Three HTLV I/II antigen (gp46, gp21) libraries are designed according to the sequence listing shown in 16A, B and C. All three HTLV-I/II antigen SSALs are synthesized as

linear peptides using the solid phase peptide synthesis strategy employing Fmoc chemistry. SSAL peptides 16A (gp46), 16B (gp46) and 16C (gp21) can be individually coated onto wells of a microplate at a concentration of 2 μg/mL. These SSAL peptides are tested with sera from patients infected with either HTLV-I or HTLV-II. SSAL HTLV-I/II peptides thus demonstrated their cost-effective feature in screening for HTLV-I/II antibodies when compared with the use of individual HTLV-I or HTLV-II peptide antigens.

Example 8

SSAL BASED UNIVERSAL FLU VACCINE FOR THE PROTECTION AGAINST DOMINANT HUMAN FLU VIRUS TYPES (e.g. H3N2, H1N1 AND H2N2)

Influenza A viruses belong to the *Orthomyxoviridae* family and cause febrile respiratory disease on a frequent and worldwide basis. High morbidity and mortality are typically associated with epidemics caused by influenza A viruses. Young children, the elderly and patients with chronic illnesses are especially susceptible to pneumonia resulting from infection. Antigenic variation occurs with exceptionally high frequency in influenza A virus. Mutations in the genes encoding the viral surface glycoproteins, namely the hemagglutinin (HA) and neuraminidase proteins (NA) are rapidly selected for by immune responses mounted against the virus. These mutations accumulate around the regions containing antigenic sites such that the mutants are capable of escaping immunity, thereby causing recurrent epidemics of respiratory disease on nearly an annual basis. Because the existence of large numbers of different virus strains and their capacity to escape immune surveillance through mutation, new vaccines structured to block infection by the current circulating virus strain(s) must be formulated on an annual basis.

Herein is described a realistic alternative to this approach by using structured synthetic antigen libraries (SSALs). The libraries are designed to mimic the major neutralizing sites present on the HA molecule, which is the dominant immunogen of this virus. Therefore, a single vaccine composed of a simple mixture of structured libraries, based upon the dominant HA antigenic sites which are capable of eliciting virus neutralizing antibodies, should eliminate and replace the need to prepare "customized" Flu virus vaccines in the conventional manner.

The first step in the design of an effective Flu vaccine is the identification of the primary amino acid sequence of the critical domain from Influenza A virus hemagglutinin (HA) surface glycoprotein capable of eliciting virus neutralizing antibodies.

A synthetic peptide representing the amino-terminus of this domain (residues 91-108), when used as a vaccine immunogen was found capable of protecting mice from virulent virus challenge (Muller, et al. 1982, Proc Natl Acad Sci USA 79: 569). Its carboxyl-terminus (residues 138-164), was also found to confer partial protection to mice (Shapira et al, 1984, Proc Natl Acad Sci USA 81: 2461). Alignment of the complete amino acid sequence of this domain, from residue 91-164, is hereby described for the H3N2 strains of Influenza virus A. The major structural feature of this domain is a loop formed by a disulfide bond between the Cysteine (Cys) residues at positions 97 and 139. The numbering of the amino acids begins with the first amino acid of the mature HA sequence (i.e. following cleavage of the signal sequence) as the first residue. The primary sequence of the structured synthetic library for the H3N2 viruses (i.e. SSAL1 Inf A-HA-H3N2) is described in Fig. 17.

Extending the amino acid sequence framework information obtained from H3N2 strains of influenza virus A to the H1N1

strains, a separate SSAL (i.e. SSAL2 Inf A-HA-H1N1) for H1N1 was similarly constructed as shown in Fig. 18A. However, there were two strains in the H1N1 family, namely WIL and FPR 834 that were found to harbor a single amino acid deletion within this domain. A third library was therefore constructed (SSAL3 Inf A-HA-H1N1) to represent the two sequences (Fig. 18B).

The "framework" information of the amino acid sequences representing the HA critical neutralizing site of two H2N2 strains was further extended, and used to construct a fourth SSAL (i.e. SSAL 4 Inf A-HA-H2N2) (Fig. 19).

Since synthetic peptides made based upon the sequences that immediately flank the loop structure present in the HA neutralizing domain were found capable of eliciting antibody that partially neutralizes virus infectivity, these flanking sequences are likely to form a conformationally dependent epitope important for eliciting a protective immune response. Three more SSALs representing a mutant form of the HA critical neutralizing domain based on the sequences of H3N2, H1N1 and H2N2 were thus constructed to mimic this "Base" structure by simply linking the amino acid sequences present on either side of the loop as shown in Fig. 20. These libraries represent deletions of the amino acid residues present between Cys residues at positions 97 and 139 in the original critical HA domain structure described above. The sequence of the deletion libraries (i.e. SSAL5 Inf A-HA-H3N2, SSAL6 Inf A-HA-H1N1, SSAL7 Inf A-HA-H2N2) was extended at the amino-terminus to include all residues represented between the first Cysteine residue of the loop structure and the next Cys residue on the amino-terminal side of the loop.

Based on the above structure design, seven SSALs are synthesized as branched peptides. These seven SSALs are tested alone and in combinations for the induction of virus-

specific antibody responses and virus-neutralizing antibody responses in laboratory animals. A response capable of neutralizing the broad-spectrum of human influenza viruses is anticipated. Following demonstration of virus neutralizing responses *in vitro*, mixtures of the these SSALs are formulated with appropriate vaccine vehicles and adjuvants and used to immunize mice. Following demonstration of an immune response to the influenza A HA SSALs in mice, the animals are challenged with the different types of human viruses to determine the breadth of protection conferred by this novel immunogen. Following mouse protection studies and demonstration of safety in laboratory animals, the SSAL-based vaccine is tested for efficacy in human volunteers.

Example 9

SSAL-BASED SYNTHETIC CHLAMYDIA VACCINE DERIVED FROM THE NEUTRALIZING DOMAIN OF THE MAJOR OUTER MEMBRANE PROTEIN (MOMP) FOR PROTECTION FROM *CHLAMYDIA* INFECTION

Chlamydia trachomatis is an obligate intracellular bacterial pathogen that is the leading cause of sexually transmitted disease (including cervicitis, epididymitis, urethritis and pelvic inflammatory disease) in industrialized countries, and the dominant pathogen associated with preventable blindness (i.e. trachoma) in developing countries. The many different isolates of this pathogen have been divided into three biovars based upon their nucleic acid composition, antigenicity and pathogenicity. Isolates that produce disease in humans come from the lymphogranuloma venereum and trachoma biovars. Within these two biovars, the different bacterial strains are further subdivided into at least 15 different serovars based upon serologic responses. These serovars are collected into three serogroups based largely upon antigenic

responses to the major outer membrane protein (MOMP): the B serogroup contains the B, Ba, D, E, L1 and L2 serovars; the intermediate group is composed of the F, G, K and L3 serovars; while, the C serogroup covers the A, C, H, I, and J serovars. The MOMP protein contains major antigenic structures that cause the production of serovar-, serogroup-, and species-specific antibody responses. Since it is the serovar- and serogroup-specific antibody responses, particularly the anti-MOMP responses that are capable of neutralizing chlamydial infectivity, MOMP has been the focus of subunit vaccine research and development. MOMP structure is conserved between serovars and consists of four variable domains separated by constant regions. The variable domains I and IV (i.e. VDI and VDIV) contain antigenic sites which elicit protective neutralizing antibody (Zhang et al. 1989, Infect. Immun. 57:636; Yuan et al. 1989, Infect. Immun. 57:1040)

Alignment of the amino acid sequence of the first variable domain (VDI) from the 15 known serovars of *C. trachomatis* has established two primary amino acid sequence patterns, designated VDI_a and VDI_b. The prototype VDI_a pattern is defined by the VDI domain from *C. trachomatis* serovar A (two alternate sequences for the VDI domain of serovar A exist, designated A_1 and A_2 above) and represents MOMP residues 81-106. Serovars A, C, H, I, J, K and L3 are representatives of the VDI_a pattern. The *C. trachomatis* serovar B VDI domain is the prototype VDI_b pattern and represents MOMP residues 81-104. Serovars B, Ba, D, E, L1 and L2 are representatives of the VDI_b pattern. The primary amino acid sequences for the SSALs for VDI_a and VDI_b are described in Figs. 21A and 21B.

Alignment of the amino acid sequences of the fourth variable domain (VDIV) from the 15 known *C. trachomatis* serovars has established two primary amino acid sequence

patterns, designated $VDIV_a$ and $VDIV_b$. The prototype $VDIV_a$ pattern is defined by the VDIV domain from *C. trachomatis* serovar A and represents MOMP residues 312-341. Serovars A, B, Ba, D, E, I, L1 and L2 are representatives of the $VDIV_a$ pattern. The *C. trachomatis* serovar C VDIV domain is the prototype $VDIV_b$ pattern and represents MOMP residues 312-342. Serovars C, F, G, H, J, K and L3 are representatives of the $VDIV_b$ pattern. The primary amino acid sequence for the structured synthetic epitope library for $VDIV_a$ and $VDIV_b$ is described in Figs. 22A and 22B.

Based on the above vaccine design, four SSALs were synthesized as branched peptides using Fmoc chemistry. Since the VDI and VDIV domains do not contain potent helper T cell epitopes one can be provided to both libraries. Addition of the helper epitope is required to potentiate antibody-producing B cell responses to the chlamydial SSALs. A T cell helper epitope, named A8, has been located between MOMP variable domains I and II (Su et al. 1990, J Exp Med 172:203). This helper epitope sequence was included at the amino-terminus of the VDI and VDIV libraries in an unmodified and non-degenerate form to elicit the T cell responses required for the production of neutralizing antibodies to the degenerate B cell domains present within the SSALs.

Following synthesis of the VDI and VDIV SSALs, they were formulated into experimental vaccines with Freund's Complete (FCA) and Incomplete Adjuvants (IFA) and used to immunize groups of five guinea pigs, at 100 µg of peptide per dose with FCA on week 0 and with IFA for booster immunizations on weeks 3 and 10, intraperitoneally. Blood was collected from the guinea pigs on weeks 0, 5, 8, and 12, processed into serum, then stored until evaluated by anti-peptide ELISA on homologous peptides. Every group of five

animals that had received Chlamydia MOMP SSAL displayed strong anti-peptide reactivity at weeks 5, 8, and 12, with ELISA reactivities of >0.5 A_{492nm} at serum dilution of 1:100,000. Serum samples from week 5 or 8 were pooled for each group, serially diluted and tested by dot immunoassay blot assay (Zhang et al. 1989) for cross-reactivity to the elementary bodies (EB) of a selection of Chlamydia trachomatis serovars representing each of the three serogroups (EBs of serovars B,Ba,D,E,F,G,A,C,H,I,J,and K). The anti-serum to the VDI_a, VDI_b, $VDIV_a$, and $VDIV_b$ SSAL vaccines were cross-reactive to all of the EBs at a 1:200 dilution. The reactivity of the VDI_a SSAL anti-serum was compared to guinea pig anti-serum made against an analogous but monovalent VDI peptide immunogen representing serovar A. The antibodies in the SSAL anti-serum bound to the EBs of each serovar approximately equally while the antibodies in the anti-VD1 A monovalent antiserum bound preferentially to the serovar A EBs with reduced reactivities for the heterogeneous EBs, in a clear demonstration of the improved immunogenicity offered by the SSAL. The SSAL guinea pig anti-sera can be further tested to demonstrate the capacity of the SSAL immunogens to induce antibodies that block *C. trachomatis* infections to all relevant serovars *in vitro;* prevent vaginal infections in mice; and prevent infections of the eye and sexually transmitted disease in primates. The VDI and VDIV libraries were initially administered by parenteral immunization. Following demonstration of antibody production specific to *C. trachomatis* elementary bodies that is capable of blocking infection *in vitro*, the SSALs can be incorporated into microparticles, or other appropriate biodegradable delivery systems, for oral, vaginal or ocular delivery. Since chlamydial infections occur at mucosal surfaces, antigen delivery to these

surfaces is considered critical for successful immunization to prevent these infections.

Example 10

SSALs FROM *Trypanosoma cruzi* ANTIGENS AS EXAMPLES OF PEPTIDE LIBRARIES DERIVED FROM REPEATS WITHIN THE SAME GENE AND USED AS DIAGNOSTIC REAGENTS FOR THE DETECTION OF *T. cruzi* INFECTION

Chagas' disease, caused by the parasite *T. cruzi,* is a chronic disease affecting about 20 million people in South and Central America. Transmission occurs through the bite of a hematophagous insect carrying infective trypomastigotes in its excreta. The acute phase of the infection is symptomless, with 90% of infected individuals progressing to the chronic phase. This phase is often characterized by myocarditis, and since no infective organisms are detectible at this stage, it is speculated that an autoimmune process is involved. An effective diagnostic test for Chagas' disease is urgently needed, both for early detection of acutely infected individuals, and for prevention of transmission by blood transfusion.

Several *T. cruzi* antigens have been found to be proteins characterized by regions of highly repeated sequence. In the few cases that have been investigated, the repeats constitute antigenic sites. For example, shed acute-phase antigen (SAPA) is a lipid-anchored family of proteins shed from *T. cruzi* during the acute phase of the infection. Some of the members of this family possess neuraminidase activity but all members contain variable numbers of a specific repeated sequence. The sequence of SEQ ID NO:27 has been found to be an immunodominant B-cell epitope (Prioli et al., 1992, Mol. Biochem. Parasitol 52:85) which is found in as many as 43 imperfect tandem repeats in SAPA proteins. This is a very good candidate for the design of a

diagnostic test because it is potentially reactive with all members of the family; in addition it offers the possibility of early detection during the acute phase of Chagas' disease.

The contiguous sequence of residues 596-1110 of one member of this family (lined up in rows of twelve residues) is shown in Fig. 23. An analogous region from another member of the SAPA family is shown in Fig. 23. The two proteins combined contain 57 imperfect repeats. An SSAL designed to consist of tandem repeats of several such partially degenerate dodecamers is also shown in Fig. 23.

Another *T. cruzi* antigen useful for design of a diagnostic test is known as cytoplasmic repetitive antigen (CRA). It has been demonstrated that the CRA antigen is recognized by all of a panel of 6 chagasic sera, and none of a panel of 9 sera from patients with infections that produce cross-reacting antibodies (Krieger et al., 1992, Am J Trop Med Hyg 46:427). By analogy with other *T. cruzi* antigens, repeated sequences are likely to constitute antigenic sites in this protein as well. Computer analysis of the sequence of the CRA gene with a program that detects potential epitopes in proteins shows that 21 imperfect repeats of a 14 residue sequence are highly likely to be antigenic. The 21 repeats are shown in Fig. 24. A diagnostic SSAL consists of numerous tandem repeats of this sequence. An SSAL based on tandem repeats of CRA sequences is also represented in Fig. 24.

By analogy with the previous example, it can also contain equal amounts of the two residues at any given variant position.

Diagnostic reagents based on the SSALs as designed used either individually or in combination would constitute a significant improvement over peptide reagents derived from a

single sequence because they have the potential to react with a broader spectrum of antibodies.

Based on the above design, two SSALs are synthesized as linear peptides using Fmoc chemistry. Both SSALs are employed as solid phase antigens and formulated into an immunoassay, both individually and in combination, for the detection of antibodies to *T. cruzi*.

Example 11

$SSAL_{DENGUE\ VIRUS}$-BASED VACCINE FOR IMMUNIZATION AGAINST VARIANTS OF DENGUE VIRUSES

Dengue viruses are members of the Flavivirus family and cause hemorrhagic fever and shock syndrome which result in significant morbidity/mortality in Asia, and Central/South America. Dengue viruses are divided principally into four major antigenic groups (Dengue subtypes 1-4), with each group subdivided into several antigenic variants. Vaccine development for Dengue virus has thus far proven difficult, since enhancing epitopes from one or more subtypes primes individuals for enhanced disease complications.

Structured Synthetic Antigen Libraries (SSAL) provide significant advantages for the design of Dengue vaccines, since variable protective neutralizing B cell antibody epitopes can be covered through this approach, along with T cell helper epitopes.

As an example, an SSAL derived from the critical neutralizing domain (aa 121-185) of the Env protein of the type-2 Dengue virus is constructed as shown in Fig. 25A, based on the primary amino acid sequence alignment from 33 antigenic variants. Three other SSALs derived from the corresponding Env regions, on the primary amino acid sequence alignments, from antigenic variants of Dengue subtypes 1, 3 and 4 can be constructed similarly.

These four SSALs (i.e. Dengue Env 1, 2, 3 and 4) can be combined with a library derived from amino acids 352-368 of

the Dengue 2 Env protein representing a T helper epitope capable of conferring T helper effects to heterologous flaviviruses, thus completing the requisite T-B requirements for immunization against variants of four subtypes of Dengue viruses.

The combined four SSALs can be formulated with appropriate vaccine vehicles and adjuvants and tested for their capacity to induce neutralizing antibodies in laboratory animals to block the replication of all relevant antigenic variants of Dengue virus *in vitro*. Following demonstration of neutralizing antibody production *in vivo*, the SSALs are then incorporated into microparticles or other appropriate vaccine delivery systems for efficacy testing in human volunteers at risk for infection by Dengue viruses.

Example 12

STREPTOCOCCAL M PROTEIN BASED VACCINE USING THE SSAL APPROACH

The efforts to date in producing an effective vaccine against Group A streptococcal infection have been thwarted by the serological diversity of the antiphagocytic and major protective antigen designated streptococcal M protein. Only type specific responses have been elicited using M-protein based vaccines. M protein sequences are characterized by a seven amino acid periodicity (Hosein, B., McCarty, M., and Fischetti, V.A., Proc. Natl. Acad. Sci. USA, 76, 3765, 1979) that allows them to assume an alpha-helical coli structure that is rare among bacterial surface molecules but is common in mammalian proteins. The extreme variability at the amino terminus has formed the basis for serological typing of M proteins, with over 100 distinct M types being identified to date.

The N-terminus of M proteins from across a number of types has been shown to produce opsonic, or antiphagocytic, antibodies. In humans and experimental animals, systemic immunity to group A streptococcal infections is due primarily to these type-specific opsonic antibodies directed against the variable amino-terminal half of M proteins. An SSAL based on an antigenic peptide structure derived from the N-terminus of the streptococcal M protein was constructed. The N-terminus of Type 5 M protein was aligned as a prototype with corresponding regions of M proteins from Types 1 and 6 in such a way as to preserve the register across types of the seven-amino acid periodicity. Despite the wide diversity of sequences across types, Leucine is completely conserved in positions 28 and 35, thus providing an anchor structure for the SSAL. Corresponding regions of M protein from Types 19 and 24 are also aligned for sequence comparison, although the extreme diversity of these two sequences has prevented us from including them into the construction of an SSAL streptococcal M.

A 35 mer SSAL beginning at the invariant C-terminal Leucine position designed according to the mathematical and chemical representation of the library (Fig. 26), is synthesized onto a heteromeric (Lysine and Alanine copolymer core) octa-branched carrier employing Fmoc chemistry.

Following synthesis of the $SSAL_{streptococcal\ M}$, this SSAL is formulated with appropriate vaccine vehicles and adjuvants and tested for its capacity to elicit protective opsonic, or anti-phagocytic, antibodies in laboratory animals.

Following demonstration of opsonic antibody production directed against multiple strains of streptococcus pyogenes, the SSAL is then incorporated into microparticles or other appropriate vaccine delivery systems for efficacy testing in human volunteers at risk for streptococcal infection.

Example 13

SSALCOLLAGEN Type I/II (Th)-BASED TREATMENT OF RHEUMATOID ARTHRITIS

Rheumatoid arthritis (RA) is a common autoimmune disease in which the joints become painful and swollen under immune system attack. The symptoms of RA can be relieved somewhat by treatment with steroids or anti-cell proliferation drugs, but these therapies have serious side effects. Recently, Trentham et al. (1993, Science 261:1727) have described a treatment of RA in human volunteers by oral tolerization with collagen.

In a related study in a mouse model of collagen-induced arthritis (Myers et al. 1993, J. Immunol. 150:4652), a peptide from collagen that was an effective competitor with collagen for MHC Class II molecules was identified. When mice were co-immunized with collagen and this MHC class II binding collagen peptide, the incidence and severity of arthritis were reduced. In another related study (Ku et al. 1993, Eur. J. Immunol. 23:591), a collagen peptide was found to prevent the development of arthritis in the rat model.

Resistance to collagen-induced arthritis in rats and rhesus monkeys was found following an immunization protocol with attenuated type-II collagen aimed at maintaining T cell epitopes, thus implying that collagen-related T cell epitopes are important in inducing a tolerized T cell state, thus preventing the development or induction of the autoimmune disease (Hart et al. 1993, Eur. J. Immunol. 23:1588).

From the above observations, we therefore adopted the use of a collagen derived T cell epitope SSAL, an approach to broaden the responder population for the treatment of patients with RA.

More specifically, a collagen T cell epitope SSAL (12mer in size) as shown in Fig. 27 can be used to immunize patients with RA with a total of from 100 to 500 μg per dose, emulsified in incomplete Freund's adjuvant at multiple sites and administered intradermally. The patients are monitored at entry, 1 month, 2 months and 3 months after immunization for various clinical parameters and symptoms. It is predicted that SSALCOLLAGEN Type I/II (Th)-based immunization will induce T cell tolerization to the collagen, thus relieving the symptoms of RA.

Example 14

$SSAL_{selectins}$-BASED TREATMENT OF INFLAMMATION

Developing effective anti-inflammatory compounds is a major goal for medical sciences. In a recent report, pertussis toxin was found to bind target cells through the carbohydrate recognition properties of two subunits, S2 and S3, which share amino acid sequence similarity with the lectin domains of the eukaryotic "selectin" family.

Selectins are a group of proteins that play an important role in promoting the reversible rolling of leukocytes at sites of tissue inflammation. To date, three selectins have been characterized and all shared a terminal lectin domain. P-selectin and E-selectin appear on activated endothelial cells and mediate adherence of leukocytes by recognition of sialylated and non-sialylated determinants. L-selectin (human leukocyte homing receptor, hLHR) is present on lymphocytes and mediates homing to lymphoid tissue that bears sialyated glycoconjugates.

More importantly, this report found that pertussis toxin S2 and S3 subunits, and synthetic peptides representing their carbohydrate recognition domains, competitively inhibit adherence of neutrophils to selectin-coated surfaces and to endothelial cells.

Based on the above observation, an SSAL representing the carbohydrate recognition domains is useful to treat inflammation. A careful examination of peptide sequences derived from multiple selectins and pertussis toxin subunits that block neutrophil adherence has allowed construction of an SSAL as shown in Fig. 28, based on selectins or peptides from prokaryotic organisms with related sequences, for the treatment of inflammation.

More specifically, a selectin SSAL is synthesized as linear peptides according to the mathematical and chemical representation shown in Fig. 28. Following the synthesis, the SSAL selectins can be formulated with an appropriate delivery system for application at site of inflammation.

Example 15

HELPER T CELL (Th) EPITOPE BASED SSALs ENHANCE ANTIBODY RESPONSE TO LHRH IN A GENETICALLY DIVERSE POPULATION

B cell responses to a foreign antigen, resulting in the production of antibodies to that antigen, are modulated by T cell responses. During antigen processing, domains or epitopes that are composed of specific linear segments of the antigen are recognized by and are presented in the context of MHC class II molecules on the surfaces of antigen presenting cells. The peptide segments presented in this fashion are recognized by $CD4^+$ T cells. These activated T cells stimulate [termed helper T cells (Th)] or suppress [termed suppressor T cell (Ts)] antibody production to B cell epitopes present on the foreign molecule in an antigen specific manner. Often, synthetic peptide antigens are poor immunogens because they do not effectively stimulate Th responses. This lack of immune reactivity can be overcome

by the addition of known Th epitopes to the synthetic antigen.

A number of Th epitopes have been characterized to date. They range in size from approximately 15-30 amino acid residues in length and often share common structural features and may contain specific landmark sequences. For example, a common feature is amphipathic helices, which are alpha-helical structures with hydrophobic amino acid residues dominating one face of the helix and with charged and polar residues dominating the surrounding faces. Th epitopes frequently contain additional primary amino acid patterns such as a Gly or charged residue followed by two to three hydrophobic residues, followed in turn by a charged or polar residue. This pattern defines what are called Rothbard sequences. Also, Th epitopes often obey the 1, 4, 5, 8 rule, where a positively charged residue is followed by hydrophobic residues at the fourth, fifth and eighth positions after the charged residue. Since all of these structures are composed of common hydrophobic, charged and polar amino acids each structure can exist simultaneously within a single Th epitope.

A further class of Th epitopes have been defined as being universal, in that they are capable of stimulating B cell responses in most members of a population that expresses diverse HLA haplotypes. A representative of this class of epitope is more appropriately described as promiscuous, because a majority, but not all, individuals expressing different HLA antigens are capable of responding to it. Most, if not all, of the promiscuous T cell epitopes contain at least one and usually more than one of the Th epitope features described above.

The design and synthesis of two SSALs that act as artificial Th epitopes and can elicit immune responses to

luteinizing hormone releasing hormone (LHRH) in all members of a genetically diverse population are described below.

The general feature of such SSALs is that the degenerate Th sequence is separated from the LHRH sequence, positions 19-28 of SEQ ID NO:34, by two Gly residues. The Gly residues act as a spacer between the T_h epitope and the LHRH sequence. The Gly spacer and the LHRH sequence are invariant in the library sequences.

The degenerate helper T cell epitope present in SSAL1 TH1LHRH is modeled after a universal epitope identified from the F protein of measles virus (Partidos et al. 1990, J Gen Virol 71: 2099). This epitope is represented by residues 288-302 of the F protein. Charged residues Glu/Asp are added at position 1 to increase the charge surrounding the hydrophobic face of the epitope. This face is made up of residues at positions 2, 5, 8, 9, 10, 13 and 16. The hydrophobic residues commonly associated with promiscuous epitopes are added at these positions. A Rothbard sequence is indicated by the residues in bold type in the prototype sequence. Sequences obeying the 1, 4, 5, 8 rule are underlined in the prototype sequence.

The degenerate helper epitope present in SSAL2 TH2LHRH is modeled after the universal epitope identified from the hepatitis B virus surface antigen (Greenstein et al. 1992, J Immunol 48: 3970). This epitope is represented by residues 19-33 of HBsAg. Positively charged Lys and Arg residues are added at positions 1, 2, 3 and 5 to increase the charge surrounding the hydrophobic face of the helical structure. The charged amino acids at position 3 also contributes the required residue to generate a sequence obeying the 1, 4, 5, 8 rule (residues underlined in the prototype sequence). Hydrophobic residues at positions 4, 6, 7, 10, 11, 13, 15 and 17 make up the hydrophobic face of an amphipathic helix.

A Rothbard sequence is indicated by the residues in bold type in the prototype sequence.

The SSALs described above and as shown in Figs. 29 A and B are therefore the first artificial members of the promiscuous class of helper epitopes. The synthetic Th: LHRH library can be used as a therapy for the treatment of androgen driven diseases including prostate cancer, prostatic hyperplasia, breast cancer and endometriosis. In addition, this same synthetic antigen can be used as a contraceptive in both males and females, with a focus on companion animals. It will also be used to prevent the syndrome commonly known as "boar taint", the offensive aroma and flavor of meat caused by the presence of testosterone in intact male food animals. The synthetic peptide libraries can be formulated in an appropriate carrier, which may include but is not limited to saline, alum, mineral or vegetable oil emulsions, microparticles, other biodegradable vehicles, and the like. Additional compounds can also be added as adjuvants to further potentiate the immune response to the SSAL formulations including, but not limited to, pluronic polymers, lipid amines, saponin, muramyl dipeptides or their analogues, cytokines, derivatives of cytokines, and lipids which may be either mixed with or covalently linked to the synthetic peptide libraries. The final formulation can be administered by immunization, oral delivery or delivery to other mucosal surfaces.

The production of structured synthetic T_h: LHRH libraries provides an ideal model for testing the efficacy of the synthetic universal helper epitopes, since LHRH is a self molecule and therefore not antigenic, i.e. antibody is not produced to unmodified LHRH when exogenously administered.

Example 16

$SSAL_{HLA\text{-}B27\ self\text{-}peptides}$ BASED TREATMENT OF HLA-B27 RELATED ARTHRITIS

The MHC molecule HLA-B27 is very strongly linked with ankylosing spondylitis and related arthropathies, including reactive arthritis. It is thought that these autoimmune diseases are caused by cytotoxic T cells reacting against an arthritogenic peptide of cellular or microbial origin, which is selectively presented by HLA-B27. The stability of HLA molecules on the surface of a cell is dependent on the binding of a peptide, and in an uninfected cell HLA molecules are stabilized by self-derived peptides. When an individual with HLA-B27 becomes infected with a microorganism which contains an arthritogenic peptide, cytotoxic T cells recognize the HLA-B27 peptide complex and attack the cell leading to autoimmune disease. An SSAL based on the sequences of naturally produced self peptides can be used to competitively occupy the peptide binding sites and prevent the presentation of the arthritogenic peptide by HLA-B27.

SSAL peptides based on the sequences of naturally processed self-peptides from HLA-B27 molecules can be used as competitive blockers to prevent presentation of arthritogenic peptides. An SSAL HLA-B27 self-peptides is synthesized as linear peptides according to the mathematical and chemical representation shown in Fig. 30. Following the synthesis, the SSAL HLA-B27 self-peptides can be formulated in an appropriate delivery system for application at the site of inflammation.

Example 17

SSAL$_{CTL\ epitope}$-BASED VACCINE THERAPY FOR PREVENTION AND TREATMENT OF DISEASE

Cytotoxic T lymphocytes (CTL) are a vital part of the immune response in many diseases, including infectious diseases such as influenza and AIDS, and in some malignant diseases such as malignant melanoma. A strong cellular response is an important component of a prophylactic vaccine and, in addition stimulation of the CTL response, can be used therapeutically to improve selected CTL responses. The stimulation of this cellular response can be achieved by the use of peptides representing target epitopes within a protein, such as the gag protein of HIV. The advantage of the SSAL$_{CTL\ epitope}$ based vaccine therapy is that sequences from a large number of related but not identical sequences in a virus can be covered. CTL epitopes that are restricted by the HLA-B27 molecule, from virus variants of HIV-1, HIV-2 and variants found from sequencing patients viruses can be incorporated into a SSAL-based vaccine. This gives a set of peptides which can stimulate a broadly-reactive CTL response and overcome possible escape mutant sequences. The SSAL$_{CTL\ epitope}$ vaccine can be encapsulated with a microparticle to be administered parenterally. In this example (see Fig. 31), the SSAL$_{CTL\ epitope}$ vaccine therapy is designed on an epitope within the gag protein of HIV that has been shown to be restricted by HLA-B27. The SSAL can be administered to HIV-infected patients with HLA-B27 to stimulate an HIV specific CTL response. Patients can be monitored at entry and each month following administration for improvement in immune function and clinical status.

Example 18

DIAGNOSTIC MARKERS AND VACCINE FOR LYME DISEASE

Lyme disease or Lyme borreliosis is an inflammatory illness in humans and other mammals with varied involvements of the skin, heart, liver, kidneys, musculoskeletal system, and the central and peripheral nervous systems (Steere. 1989, N Engl J Med 321:586). The etiological agents are a group of spirochetes which are collectively known as *Borrelia burgdorferi*, and which are transmitted by ticks of the genus *Ixodes*. The disease is of global distribution and is the most prevalent tick-borne disease in the U.S. Prevention is by tick-control methods, since there is no effective vaccine available.

The organisms responsible for Lyme disease have been subdivided by recent phylogenetic analysis into at least three genospecies referred to as *B. burgdorferi* sensu strictu, *Borrelia garinii*, and *Borrelia afzelii* (formerly known as VS461). *B. burgdorferi* sensu strictu is common in Europe and North America. It accounts for over 90% of isolates in the United States. It tends to be correlated with arthritis. *Borrelia garinii* is found in Europe and tends to be associated with neuroborreliosis. *B. afzelii* isolates have been found in Europe and Japan and are linked to acrodermatitis chronica atrophicans (Cann et al, 1992 Lancet 339:1598). Antigenic variation can be correlated to these subgroups by serotypic and sequence differences in major outer surface proteins OspA and OspB. This variation occurs at both inter- and intra-species levels. The influence of antigenic variation and strain specificity can be observed in the limited cross-protection between isolates in animal vaccine trials with Osp subunit antigens (Marconi

et al. 1993 Infect Immun 61:2611) and in a lack of natural cross-immunity in humans.

Recombinant proteins corresponding to variable major outer surface proteins (Osp) of B. burgdorferii, OspA and OspB have been proven in mouse models to elicit protective neutralizing antibodies and to have potential as protective vaccines (Telford et al. 1993 J Exp Med 178:755). OspA is the more conserved of the two among North American isolates but was found to display considerable variability (Wallich et al. 1992 Infect Immun 60: 4856). OspA and OspB vaccines formulated to correspond to single strains do not elicit strongly cross-protective immunity. SSALs can provide significant advantages for the design of Lyme disease vaccine, since variable protective epitopes can be effectively included for wide efficacy.

OspA antigenicity has been extensively analyzed. Studies with monoclonal antibodies and patient antibodies have shown that the C-terminal half, residues 108 to 273, is exposed to the human immune system (Schubach et al. 1991 Infect Immun 59:1911-1915). This half of OspA is highly variable between strains and bears one or more crucial non-contiguous antigenic determinants between residues 133 to 273 which are responsible for strain-specific protective immunity. Class II T cell determinants have been identified at residues 221-235, 258-273, and 248 to 263 and B cell epitopes are recognized by human serum samples on the terminal 60 residues, 211-273 (Shanafelt et al. 1992 J Immunol 148:218).

Two useful SSALs depicted in Figs. 32A and B can be constructed as OspA SSAL immunogens, based on the primary amino acid sequence alignment from 12 antigenically variant isolates and antigenic mapping. They are useful for the detection of antibodies to *Borrelia* in sera and body fluids when formulated into immunoassays, and they are key

components of a protective vaccine for humans and animals when incorporated into microparticles or other appropriate vaccine delivery system.

OspC is another heterogeous outer surface protein common to the three genospecies of the *Borrelia burgdorferi* group. OspC elicits a prominent antibody response early in infection making it a useful diagnostic marker and studies with animal models have shown it to be a candidate for an effective vaccine (Jauris-Heipke et al, 1993). It is somewhat more variable in sequence and antigenicity than OspA, a heterogeneity that suggests that it can be effectively applied to diagnostics and vaccines as an SSAL antigen. An SSAL (Fig. 33) useful for those applications can be constructed based on an alignment of primary amino acid sequences from OspC of 15 *Borrelia* isolates. These sequences were published: 13 in the NCBI CD/ROM database "Entrez" as NCBI seq ID nos. 311392, 311394, 313272, 313274, 313276, 313278, 313280, 434658, 434662, 434664, 434666, 443725, and 495736; and two in GenBank as accession nos. U08284 and X62162. The SSAL immunogen as shown in Fig. 33 is incorporated into vaccine and diagnostic compositions.

A central region of the primary sequence for the flagellin protein of *B. burgdorferi* comprises a cluster of antigenic determinants that cause the production of antibodies specifically associated with Lyme disease. Serological detection and measurement of these antibodies is useful for the diagnosis of Lyme disease and for the prognosis of clinical outcome (Schneider et al, 1992 <u>Infect Immun</u> <u>60</u>:316). This segment is encompassed by a region spanning flagellin residues 175-235 that is described by the SSAL of Fig. 34. This SSAL was designed in accordance with amino acid sequences deduced for 12 isolates from North America and Europe (Collins and Peltz, 1991 <u>Infect Immun</u>

59:514; Picken, 1992 J Clin Microbiol 30:99; Jauris-Heipke et al, 1993 Med Microbiol Immunol 182:37). It includes a conserved framework sequence that is species-specific and common to the isolates of *B. burgdorferi*, and ten positions within the framework sequence that vary among the isolates. Though flagellin sequence is more conserved than OspA sequence for the relevant genospecies of *Borrelia*, the pattern of isolate to isolate variation for flagellin does display some correlation with the heterogeneity of OspA (Jauris-Heipke et al, 1993; Picken, 1992).

Antibodies to regions of flagellin outside of the segment described by the SSAL of Fig. 34 share homologies and cross-reactivities to the flagellins of other bacteria. Thereby, a reduced specificity for the detection of Lyme disease is imparted to immunoassays which are sensitized with flagellin antigens that include the cross-reactive domains (Schneider et al, 1992). Conversely, the isolate-specific variation of the species-specific central domain suggests reduced antigenicity for immunoassays which are sensitized to flagellin by single-strain central region peptides. Thus, a specific and comprehensive SSAL such as that of Fig. 34 is advantageously employed in immunoassays for the diagnosis and prognosis of Lyme disease. A peptide composition comprised of the SSAL depicted in Fig. 34 is useful for the detection of antibodies to the *Borrelia burgdorferi* group of bacteria in sera and body fluids in North America and Europe, and contributes to the diagnosis and prognosis of Lyme disease.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: United Biomedical, Inc.

(ii) TITLE OF INVENTION: Structured Synthetic Antigen Libraries (SSAL) for Diagnostics, Vaccines and Therapeutics

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Maria C.H. Lin, Esq.
(B) STREET: 345 Park Avenue
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10154

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WordPerfect 5.1

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 26-OCT-1994
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/143,412
(B) FILING DATE: 26-OCT-1993

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Lin, Maria C.H.
(B) REGISTRATION NUMBER: 29,323
(C) REFERENCE/DOCKET NUMBER: 1151-4120PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 212-758-4800
(B) TELEFAX: 212-751-6849

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 47 amino acids
    (B) TYPE: amino acid
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 1
    (D) OTHER INFORMATION: /note= "S8;N2"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 2
    (D) OTHER INFORMATION: /note= "G8;D1;Q1"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 3
    (D) OTHER INFORMATION: /note= "K6;R4"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 4
    (D) OTHER INFORMATION: /note= "P8;V1"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 5
    (D) OTHER INFORMATION: /note= "A8;V2"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 6
    (D) OTHER INFORMATION: /note= "I6;V4"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 7
    (D) OTHER INFORMATION: /note= "I6;V2;A2"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 10
    (D) OTHER INFORMATION: /note= "R8;K2"

(ix) FEATURE:
    (A) NAME/KEY: Modified-site
    (B) LOCATION: 12

(D) OTHER INFORMATION: /note= "V9;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "R3;Q5;E2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "E8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "S4;A6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "Q4;S6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "H8;K1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "L8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "P8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "Y8;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "I9;F1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "Q8;E2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "M8;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "M4;Q4;R2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "L8;M1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "Q8;M2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "F8;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "Q8;S2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "A8;I2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "L8;Q2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu
1               5                   10                  15

Phe Asp Glu Met Glu Glu Cys Ser Gln His Leu Pro Tyr Ile Glu Gln
            20                  25                  30

Gly Met Met Leu Ala Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "P9;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "K3;R7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "N3;T7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "V9;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 48
(D) OTHER INFORMATION: /note= "T9;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "S9;W1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Ser Thr Asn Pro Lys Pro Gln Lys Lys Asn Lys Arg Asn Thr Asn Arg
1               5                   10                  15

Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly
            20                  25                  30

Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala Thr
        35                  40                  45
```

Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg
50 55 60

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "K9;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "L6;I4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "E8;A1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "V4;A3;T1;F1;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "K9;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "K8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "V4;S3;T1;R2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32

(D) OTHER INFORMATION: /note= "A5;G5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "L8;M2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "I6;L3;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "N9;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 51
(D) OTHER INFORMATION: /note= "T9;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "S7;I1;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 54
(D) OTHER INFORMATION: /note= "D9;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 59
(D) OTHER INFORMATION: /note= "A9;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 67
(D) OTHER INFORMATION: /note= "Y5;F5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 79
(D) OTHER INFORMATION: /note= "T8;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 80
(D) OTHER INFORMATION: /note= "C8;A2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile Phe
1               5                   10                  15

Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Ala
            20                  25                  30

Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser Val
        35                  40                  45

Ile Pro Thr Ser Gly Asp Val Val Val Val Ala Thr Asp Ala Leu Met
    50                  55                  60

Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr Cys
65                  70                  75                  80
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "R9;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "P8;L1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "V5;L4;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "T5;S4;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "K4;D4;R2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "D9;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "E5;V4;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "P9;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "V8;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "H8;L1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "P8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "P9;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "P5;T3;V1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "K8;P1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33

(D) OTHER INFORMATION: /note= "S4;A4;K1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "P8;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "V5;I4;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "K7;R3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "T9;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "K4;R6"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro
1               5                   10                  15

Asp Tyr Glu Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Lys
            20                  25                  30

Ser Pro Pro Val Pro Pro Pro Arg Lys Lys Arg Thr
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "E8;D2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "V8;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "A8;E1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "T9;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "A6;T4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "G5;K3;A2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Glu Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr Asp
1 5 10 15

Pro Ser His Ile Thr Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg Gly
20 25 30

Ser Pro Pro Ser
35

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "L12;I2;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "V7;T4;I5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "A12;V4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "Y12;V2;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "Q5;E11"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "R11;H3;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "S5;V7;I4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "T3;S13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "G12;S2;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "L5;I7;S2;G2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "H12;Y2;M2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "V14;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "P5;S9;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "S10;A2;N2;D2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "I15;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "V12;T4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "Y12;W4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "E12;Q4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "A12;L4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "A11;H1;T2;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35

(D) OTHER INFORMATION: /note= "D13;A2;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "A9;L2;M4;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "I12;V4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "L9;M7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "T11;L2;V2;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "V12;E4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 48
(D) OTHER INFORMATION: /note= "R12;N2;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 49
(D) OTHER INFORMATION: /note= "E12;D2;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "G9;S2;N4;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 51
(D) OTHER INFORMATION: /note= "N14;G2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 53
(D) OTHER INFORMATION: /note= "S14;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "A4;S5;F2;T4;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 54
(D) OTHER INFORMATION: /note= "R15;H1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Cys Leu Thr Val Pro Ala Ser Ala Tyr Gln Val Arg Asn Ser Thr Gly
1 5 10 15

Leu Tyr His Val Thr Asn Asp Cys Pro Asn Ser Ser Ile Val Tyr Glu
20 25 30

Ala Ala Asp Ala Ile Leu His Thr Pro Gly Cys Val Pro Cys Val Arg
35 40 45

Glu Gly Asn Ala Ser Arg Cys
50 55

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "V12;E4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "R12;K2;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "E12;V2;D2"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "G9;S2;N4;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "N14;G2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "A4;S5;F2;T4;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "S14;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "R15;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "V12;I4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "A12;Q2;P2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "M3;V6;L7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "T14;S2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "T12;N4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "V9;L7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "T5;A7;V4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "R12;Q2;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "D5;N7;Q2;H2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "G5;A2;S2;V2;P2;R2;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "K5;S5;T2;G4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "L5;V3;I4;A4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "P12;L4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "A3;T13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "T11;Q2;H1;A1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29

(D) OTHER INFORMATION: /note= "Q5;T7;G2;N1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "L10;I6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "R12;T4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "I7;V9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "L12;M4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "L12;I2;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "G12;M4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "S7;A8;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "T9;A7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "L9;F5;V1;A1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp Val Ala Met Thr Pro

```
1               5                   10                  15
Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Ala Thr Gln Leu Arg Arg
            20                  25                  30
His Ile Asp Leu Leu Val Gly Ser Ala Thr Leu Cys
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "T12;I3;M1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "F12;V3;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "R12;Q3;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "R13;H3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "H13;Y3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "W7;E5;V2;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10

(D) OTHER INFORMATION: /note= "T12;F4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "T7;V7;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "G3;D11;E2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "P14;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "H15;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "I9;L1;V6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "T10;S6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "D15;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "M14;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "N15;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "T8;A6;L2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Gly Cys Asn Cys
1   5   10   15

Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp Asp Met
            20   25   30

Met Met Asn Trp Ser Pro Thr Thr
        35   40

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "A5;G5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "E3;H2;D2;S1;T1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "H7;I2;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "V7;S1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "T6;S3;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "G9;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "S2;Q1;R2;A2;N1;T1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "H2;R4;S2;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "T6;S2;N1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "V2;T6;M1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "S2;A2;N2;Q2;H1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "G4;S3;R2;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "F2;L8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "V7;A2;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "S7;G3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "L5;M3;W2"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "A3;S3;T4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "P3;Q3;S2;T1;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "A6;P4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "K4;S5;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "N4;K6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "Q9;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "I8;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "A4;Q2;V3;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "G4;A6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "L5;F5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "V4;I5;L1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Asp Ala Glu Thr His Val Thr Gly Gly Ser Ala Gly His Thr Val
1 5 10 15

Ser Gly Phe Val Ser Leu Leu Ala Pro Gly Ala Lys Gln Asn Val Gln
20 25 30

Leu Ile Asn
35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "Q8;I1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "P8;T1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "I8;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "S5;T3;E1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "Y9;H1"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "A6;D1;T1;E2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "N5;E2;M1;T1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "*8;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "*8;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "G5;S1;P1;A1;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "S6;E1;D1;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "G5;R1;S1;I1;D1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "P4;L1;S2;Q1;G1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "Q7;E1;M2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "P8;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "K4;R5;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "P8;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "G9;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "I9;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "P9;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "K5;S4;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "S6;E1;Q2;T1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Xaa Xaa Gly Ser Gly Pro
1               5                   10                  15

Asp Gln Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys Gly Ile
            20                  25                  30

Val Pro Ala Lys Ser Val Cys Gly Pro Val Tyr Cys
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "E12;K4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "S11;P2;A2;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "E9;R2;K2;Q1;T1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "N15;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "T15;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "P14;L1;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "N12;S1;Y3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "N13;Y1;K1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "N15;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14

(D) OTHER INFORMATION: /note= "T13;K1;V1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "R14;K1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "K9;R3;T1;E1;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "S11;R2;G3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "I10;V3;L1;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "H8;P4;R2;G1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "I12;L2;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "P14;W1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "Q9;R6;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "A11;T3;V1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "F10;W2;L3;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "Y12;F2;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "A6;T7;R2;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "T14;R1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "G13;K1;R2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "D7;E3;N3;Q1;Y1;R1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys
1               5                   10                  15

Ser Ile His Ile Gly Pro Gly Gln Ala Phe Tyr Ala Thr Gly Asp Met
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "I9;V38;L3"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "L49;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "V41;L5;I4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "R47;S1;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "Y47;F1;L2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "L49;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "K27;Q6;G1;R16"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "D47;E1;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "Q49;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "L48;R1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "L49;M1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "I42;F2;M2;G1;L3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "W49;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "G49;W1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "C49;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "S48;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "K49;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "L43;H4;I2;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "I47;V3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "T46;Y3;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "A22;T20;N5;F1;S2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31

(D) OTHER INFORMATION: /note= "P48;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "A26;T8;S14;R1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "S48;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "G46;E1;R1;N1;S1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly
1               5                   10                  15

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp
            20                  25                  30

Asn Ala Ser
        35
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "T19;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "Q1;K18;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "D19;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "R2;Q16;L1;X1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "L19;X1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "S14;A6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "T18;S1;X1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "W19;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "V7;A1;P12"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "D10;E4;A6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "S11;T8;N1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Arg Val Thr Ala Ile Glu Lys Tyr Leu Gln Asp Gln Ala Arg Leu Asn
1               5                   10                  15

Ser Trp Gly Cys Ala Phe Arg Gln Val Cys His Thr Thr Val Pro Trp
            20                  25                  30
```

Val Asn Asp Ser
35

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "L1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "N1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "T1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "S1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "L1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "A1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "L1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17

(D) OTHER INFORMATION: /note= "P1;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "H1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "W1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "D1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "I1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "E1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "I1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "P1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "K1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "S1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "L1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "T1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "L1;F1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "V1;I1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Leu Asn Thr Glu Pro Ser Gln Leu Pro Pro Thr Ala Pro Pro Leu Leu
1               5                   10                  15

Pro His Ser Trp Leu Asp His Ile Leu Glu Pro Ser Ile Pro Trp Lys
            20                  25                  30

Ser Lys Leu Leu Thr Leu Val Gln Leu Thr Leu Gln Ser
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "F1;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "D1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "Q1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "I1;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "V1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "S1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "S1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "P1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "L1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "S1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "T1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "G1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "S1;T1"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "S1;R1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Cys Phe Asp Pro Gln Ile Gln Ala Ile Val Ser Ser Pro Cys His Asn
1 5 10 15

Ser Leu Ile Leu Pro Pro Ser Leu Ser Pro Val Pro Thr Leu Gly Ser
20 25 30

Arg Ser Arg Arg Ala
35

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "Q1;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "K1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "L1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "K1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "I1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "N1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "L1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "R1;C1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 54
(D) OTHER INFORMATION: /note= "P1;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 57
(D) OTHER INFORMATION: /note= "T1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 59
(D) OTHER INFORMATION: /note= "S1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 62
(D) OTHER INFORMATION: /note= "P1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 63
(D) OTHER INFORMATION: /note= "I1;V1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Val Asp Lys Asp Ile Ser Gln Leu Thr Gln Ala Ile Val Lys His His
1               5                   10                  15

Lys His Leu Leu Lys Ile Ala Asn Tyr Ala Ala Gln His Arg Arg Gly
            20                  25                  30

Leu Asp Leu Leu Phe Trp Glu Gln Gly Gly Leu Cys Lys Ala Leu Gln
        35                  40                  45
```

Glu Gln Cys Arg Phe Pro His Ile Thr His Ser His Val Pro Ile Leu
50 55 60

Gln Glu Arg Pro Pro Leu Glu
65 70

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "S95;G5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "K82;N18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "Y86;F14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "S95;Y5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "F95;C5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "N50;T45;S5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "G90;N5;D5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "T60;N40"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "W90;L10"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "G95;E5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "N77;S23"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "S60;Y30;S10"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "G90;E10"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 53
(D) OTHER INFORMATION: /note= "P72;S27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 54
(D) OTHER INFORMATION: /note= "D46;V18;A18;G18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 55
(D) OTHER INFORMATION: /note= "N82;S14;K4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 56
(D) OTHER INFORMATION: /note= "G68;S32"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 65
(D) OTHER INFORMATION: /note= "Y55;T45"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 66
(D) OTHER INFORMATION: /note= "K82;E18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 68
(D) OTHER INFORMATION: /note= "G64;E36"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 69
(D) OTHER INFORMATION: /note= "S86;Y14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 70
(D) OTHER INFORMATION: /note= "T72;K27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 73
(D) OTHER INFORMATION: /note= "V86;A14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 74
(D) OTHER INFORMATION: /note= "L91;Q9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "S91;N9"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Ser Lys Ala Tyr Ser Ser Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1               5                   10                  15

Ser Leu Arg Ser Leu Val Ala Ser Ser Gly Thr Leu Glu Phe Ile Asn
            20                  25                  30

Glu Gly Phe Thr Trp Thr Gly Val Thr Gln Asn Gly Gly Ser Ser Ala
        35                  40                  45
```

Cys Lys Arg Gly Pro Asp Asn Gly Phe Phe Ser Arg Leu Asn Trp Leu
50 55 60

Tyr Lys Ser Gly Ser Thr Tyr Pro Val Leu
65 70

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "T85;I15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "Y85;D15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "A85;I15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "S85;F15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "W84;N8;T8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "P84;W8;L8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "N62;K23;T15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 43
 (D) OTHER INFORMATION: /note= "H85;G15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 44
 (D) OTHER INFORMATION: /note= "T69;N31"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 45
 (D) OTHER INFORMATION: /note= "V70;T15;Q15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 47
 (D) OTHER INFORMATION: /note= "K62;R23;N15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 49
 (D) OTHER INFORMATION: /note= "V85;G15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 50
 (D) OTHER INFORMATION: /note= "T85;S15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 51
 (D) OTHER INFORMATION: /note= "A84;S8;N50"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 52
 (D) OTHER INFORMATION: /note= "S69;A31"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 54
 (D) OTHER INFORMATION: /note= "S85;K15"

(ix) FEATURE:
 (A) NAME/KEY: Modified-site
 (B) LOCATION: 55
 (D) OTHER INFORMATION: /note= "H85;R15"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 56
(D) OTHER INFORMATION: /note= "K55;A15;N15;G15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 57
(D) OTHER INFORMATION: /note= "G84;P8;R8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 58
(D) OTHER INFORMATION: /note= "K70;D15;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 60
(D) OTHER INFORMATION: /note= "S85;G15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 62
(D) OTHER INFORMATION: /note= "Y85;F15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 63
(D) OTHER INFORMATION: /note= "R85;S15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 64
(D) OTHER INFORMATION: /note= "N85;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 66
(D) OTHER INFORMATION: /note= "L85;N15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 69
(D) OTHER INFORMATION: /note= "T85;Y15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 70
(D) OTHER INFORMATION: /note= "E70;G15;K15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 71
(D) OTHER INFORMATION: /note= "K85;S15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 72
(D) OTHER INFORMATION: /note= "N77;G15;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 73
(D) OTHER INFORMATION: /note= "G77;S15;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 74
(D) OTHER INFORMATION: /note= "L62;S23;T15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 77
(D) OTHER INFORMATION: /note= "N70;K15;V15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 78
(D) OTHER INFORMATION: /note= "L92;Q8"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Pro Asn Ser Glu Asn Gly Thr Cys Tyr Pro Gly Tyr Phe Ala Asp Tyr
1               5                   10                  15

Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe
            20                  25                  30

Glu Ile Phe Pro Lys Ser Ser Trp Pro Asn His Thr Val Thr Lys Gly
        35                  40                  45

Val Thr Ala Ser Cys Ser His Lys Gly Lys Ser Ser Phe Tyr Arg Asn
    50                  55                  60

Leu Leu Trp Leu Thr Glu Lys Asn Gly Leu Tyr Pro Asn Leu
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "D1;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "G1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "S1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "R1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 57
(D) OTHER INFORMATION: /note= "N1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 70
(D) OTHER INFORMATION: /note= "E1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 72
(D) OTHER INFORMATION: /note= "S1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 73
(D) OTHER INFORMATION: /note= "D1;N1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Glu Asn Pro Arg Asp Gly Leu Cys Tyr Pro Gly Ser Phe Asn Asp Tyr
1               5                   10                  15

Glu Glu Leu Lys His Leu Leu Ser Ser Val Lys His Phe Glu Lys Val
            20                  25                  30
```

Lys Ile Leu Pro Lys Asp Arg Trp Thr Gln His Thr Thr Thr Gly Gly
35 40 45

Ser Arg Ala Cys Ala Val Ser Gly Asn Pro Ser Phe Phe Arg Asn Met
50 55 60

Val Trp Leu Thr Lys Glu Gly Ser Asp Tyr Pro Val Ala
65 70 75

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "G64;V36"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "N77;D23"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "T59;K41"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "V95;I5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "S95;G5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "K82;N18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "Y86;F14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "S90;N10"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "G90;E10"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "P73;S27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "D46;V18;A18;G18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "N82;S14;K4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "G68;S32"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "Y55;T45"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "K82;E18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "G64;E36"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42

(D) OTHER INFORMATION: /note= "S86;Y14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "T73;K27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 46
(D) OTHER INFORMATION: /note= "V86;A14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "L90;Q10"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Asp Gly Phe Gln Asn Glu Thr Trp Asp Leu Phe Val Glu Arg Ser Lys
1               5                   10                  15

Ala Tyr Ser Ser Cys Cys Lys Arg Gly Pro Asp Asn Gly Phe Phe Ser
            20                  25                  30

Arg Leu Asn Trp Leu Tyr Lys Ser Gly Ser Thr Tyr Pro Val Leu
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "E85;D15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "S85;P15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "F77;L15;V8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "S85;P15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "K70;V15;Q15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "K70;E15;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "S92;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "A85;V15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "T85;I15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "S85;K15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "S85;K15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "H85;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28

(D) OTHER INFORMATION: /note= "K55;A15;N15;G15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "G84;P8;R8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "K70;D15;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "S85;G15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "Y85;F15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "R85;S15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "N85;R15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "L85;N15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "T85;Y15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "E70;G15;K15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "K85;S15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "N77;G15;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "G77;S15;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 46
(D) OTHER INFORMATION: /note= "L62;S23;T15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 49
(D) OTHER INFORMATION: /note= "N70;K15;V15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "L92;Q8"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Glu Ser Leu Phe Ser Lys Lys Ser Trp Ser Tyr Ile Ala Glu Thr Pro
1               5                   10                  15

Asn Ser Glu Asn Gly Thr Cys Cys Ser Ser His Lys Gly Lys Ser Ser
            20                  25                  30

Phe Tyr Arg Asn Leu Leu Trp Leu Thr Glu Lys Asn Gly Leu Tyr Pro
        35                  40                  45

Asn Leu
    50
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "S1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "D1;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "G1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "N1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "E1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "S1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "D1;N1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Asp Arg Leu Leu Ser Val Pro Glu Trp Ser Tyr Ile Met Glu Lys Glu
1               5                   10                  15

Asn Pro Arg Asp Gly Leu Cys Cys Ala Val Ser Gly Asn Pro Ser Phe
            20                  25                  30

Phe Arg Asn Met Val Trp Leu Thr Lys Glu Gly Ser Asp Tyr Pro Val
        35                  40                  45

Ala
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "A7;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "S5;N1;K1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "V6;A1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "A7;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "G7;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "E3;Q4;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "K2;N6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "V2;T5;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22

(D) OTHER INFORMATION: /note= "A1;T6;V1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Glu Phe Gln Met Gly Ala Ala Pro Thr Thr Ser Asp Val Ala Gly Leu
1               5                   10                  15

Glu Lys Asp Pro Val Ala Asn Val Ala Arg
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "A5;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "T3;A2;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "T4;D1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "A5;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "V1;A3;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "S5;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "L4;C2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Glu Phe Gln Met Gly Ala Lys Pro Thr Thr Thr Thr Gly Asn Ala Val
1 5 10 15

Ala Pro Ser Thr Leu Thr Ala Arg
20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "L2;S6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "K1;T4;E3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "P1;A4;T3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "V2;I6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "L2;F6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "T4;V4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "K2;A6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "T2;D5;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "V2;K6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "S2;T3;A3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "S6;G1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "N2;G6"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "E3;Q5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "A2;G6"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Leu Ala Lys Pro Val Leu Asp Thr Thr Thr Leu Asn Pro Thr Ile Ala
1               5                   10                  15

Gly Lys Gly Thr Val Val Ser Ser Ala Glu Asn Glu Leu Ala
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "A6;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "E5;T1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "A6;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "I5;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "L6;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "V6;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "A6;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "K5;C2"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "S5;T1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "V6;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "S2;A4;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "A3;S4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "G5;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "T2;S5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "D3;E4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "N5;G2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "E3;D2;Q2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "L5;I2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "A5;S2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
Leu Ala Glu Ala Ile Leu Asp Val Thr Thr Leu Asn Pro Thr Ile Ala
1               5                   10                  15

Gly Lys Gly Ser Val Val Ser Ala Gly Thr Asp Asn Glu Leu Ala
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "S55;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "S37;G20"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "T56;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "A33;V24"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Asp Ser Ser Ala His Ser Thr Pro Ser Thr Pro Ala
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "V19;A2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "A19;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "A19;T2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "R3;K18"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "T15;M4;A2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
Val Ala Glu Ala Glu Lys Gln Arg Ala Ala Glu Ala Thr Lys
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2

(D) OTHER INFORMATION: /note= "K32;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "N31;S2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "K3;E30"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "V16;I16;F1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "Q30;L2;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "I29;V4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "I28;V5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "H22;N10;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "K32;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "I27;L1;V5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "K30;V3"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "I24;V9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "I30;M1;T1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 61
(D) OTHER INFORMATION: /note= "V32;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 64
(D) OTHER INFORMATION: /note= "E31;D1;Q1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Cys Lys Lys Asn Met Lys Gly Lys Val Val Gln Pro Glu Asn Leu Glu
1               5                   10                  15

Tyr Thr Ile Val Ile Thr Pro His Ser Gly Glu Glu His Ala Val Gly
            20                  25                  30

Asn Asp Thr Gly Lys His Gly Lys Glu Ile Lys Ile Thr Pro Gln Ser
        35                  40                  45

Ser Ile Thr Glu Ala Glu Leu Thr Gly Tyr Gly Thr Val Thr Met Glu
    50                  55                  60

Cys
65
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1

(D) OTHER INFORMATION: /note= "I32;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "N32;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "V31;A1;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "S32;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "N32;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "I31;V2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "E32;D1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
Ile Thr Val Asn Pro Ile Val Thr Glu Lys Asp Ser Pro Val Asn Ile
1               5                   10                  15

Glu
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "T1;V1;*"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "V1;F1;*"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "T1;N1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "R2;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "G2;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "T2;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "I1;N1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "S1;P1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "D1;R1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "P2;E1"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "Q1;V1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "R1;I1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "A2;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "K1;D1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "E2;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "A2;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "L2;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "D1;N2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= "K2;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "Y2;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "E1;A1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "L1;I1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24
(D) OTHER INFORMATION: /note= "E2;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "H1;I1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "D1;R1;M1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "K1;R1;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "T1;H1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "K1;E1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "E1;K1;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "G1;D1;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36

(D) OTHER INFORMATION: /note= "K2;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "T2;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "E2;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "N2;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "E2;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "G1;N1;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "L2;A1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Arg Thr Val Thr Arg Gly Thr Ile Ser Asp Pro Gln Arg Ala Lys Glu
1               5                   10                  15

Ala Leu Asp Lys Tyr Glu Leu Glu Asn His Asp Leu Lys Thr Lys Asn
            20                  25                  30

Glu Gly Leu Lys Thr Glu Asn Glu Gly Leu
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "E3;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "A3;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "Q3;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "P3;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "E3;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "P3;S1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Pro Glu Gly Ala Gln Gly Pro Arg Gly Glu Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "E1;A1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "E2;D1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "N2;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "S1;K2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "I1;T1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "S1;P2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "Y2;F1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "S2;Y1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Lys Glu Glu Ile Glu Tyr Leu Asn Ser Ile Leu Ser Tyr Ser
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "D0.50;E0.50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "L0.25;I0.25;V0.25;F0.25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "E0.50;D0.50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "L0.25;I0.25;V0.25;F0.25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "K0.50;R0.50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "K50;R50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12

(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "D50;E50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```
Asp Leu Glu Leu Lys Gly Leu Leu Leu His Lys Leu Asp Gly Leu Gly
1               5                   10                  15

Gly Glu His Trp Ser Tyr Gly Leu Arg Pro Gly
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "K50;R50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "K50;R50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "K50;R50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "F34;K33;R33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "K50;R50"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "Q20;L20;I20;F20;V20"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "L25;I25;V25;F25"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
Lys Lys Lys Leu Phe Leu Leu Thr Lys Leu Leu Thr Leu Pro Gln Ser
1               5                   10                  15
```

Leu Gly Gly Glu His Trp Ser Tyr Gly Leu Arg Pro Gly
20 25

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "R5;G1;F1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "Y2;I2;V1;W1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "Q1;K3;L1;D1;N1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "K2;E3;P1;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "S1;I2;V1;A1;P1;L1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "T2;V2;I2;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "E1;K2;D1;V1;L1;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9

(D) OTHER INFORMATION: /note= "L1;K3;R2;A1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Arg Arg Tyr Gln Lys Ser Thr Glu Leu
1 5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "K8;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "W8;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "I8;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "I5;V3;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "L5;I2;M2"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "G8;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9

(D) OTHER INFORMATION: /note= "N8;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "K8;R1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "T76;K16;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "L84;V16"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "A58;T42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "A67;N25;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "D50;E42;G8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9

(D) OTHER INFORMATION: /note= "T67;V25;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "T92;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "L92;V8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "V34;E33;K33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "V92;I8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "K75;T25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "T75;V25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "S92;K8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "K92;R8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "N50;E34;H16"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "S42;A25;L17;P8;D8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "K92;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "S92;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "E92;K8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "V75;I25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "T67;S33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "V92;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "E50;A42;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "N75;D25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "T73;S27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "T58;S42"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "T58;S34;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "Q58;A34;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 42
(D) OTHER INFORMATION: /note= "A92;G8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "T92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 47
(D) OTHER INFORMATION: /note= "G58;A42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 48
(D) OTHER INFORMATION: /note= "A58;K26;S8;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "D58;N34;Q8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 51
(D) OTHER INFORMATION: /note= "S84;A8;E8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "K42;G42;S8;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 53
(D) OTHER INFORMATION: /note= "T92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "D73;N27"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Gly Thr Leu Ala Ala Asp Gly Lys Thr Thr Leu Val Val Lys Glu Gly
1 5 10 15

Thr Val Thr Leu Ser Lys Asn Ile Ser Lys Ser Gly Glu Val Thr Val
20 25 30

Glu Leu Asn Asp Thr Asp Thr Thr Gln Ala Thr Lys Lys Thr Gly Ala
35 40 45

Trp Asp Ser Lys Thr Ser Thr Leu Thr Ile
50 55

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "N75;D25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "T73;S27"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "D73;N23"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "T58;S42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "T58;S33;A9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "Q58;A33;S9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "A92;G8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "T92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "G58;A42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "A58;K25;S8;N9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "D58;N33;Q9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "S83;A8;E9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "K42;G42;S8;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "T92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 24

(D) OTHER INFORMATION: /note= "S58;T42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "V92;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "S92;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "K75;Q17;R8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "K75;T25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "D42;N25;Q25;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 33
(D) OTHER INFORMATION: /note= "L92;I8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "T92;L8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "K92;T8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "E58;Q33;N9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "D58;N33;G9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "V92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "K58;Q33;N9"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 49
(D) OTHER INFORMATION: /note= "A75;N25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "N58;K42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 56
(D) OTHER INFORMATION: /note= "S42;T33;K17;N8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 58
(D) OTHER INFORMATION: /note= "V92;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 61
(D) OTHER INFORMATION: /note= "K58;T42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 62
(D) OTHER INFORMATION: /note= "T58;K42"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 64
(D) OTHER INFORMATION: /note= "D83;K17"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 66
(D) OTHER INFORMATION: /note= "L67;I33"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 68
(D) OTHER INFORMATION: /note= "N75;D17;A8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 71
(D) OTHER INFORMATION: /note= "R25;K75"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Asn Asp Thr Asp Thr Thr Gln Ala Thr Lys Lys Thr Gly Ala Trp Asp
1               5                   10                  15

Ser Lys Thr Ser Thr Leu Ile Ser Val Asn Ser Lys Lys Thr Lys Asp
            20                  25                  30

Leu Val Phe Thr Lys Glu Asp Thr Ile Thr Val Gln Lys Tyr Asp Ser
        35                  40                  45

Ala Gly Thr Asn Leu Glu Gly Ser Ala Val Glu Ile Lys Thr Leu Asp
    50                  55                  60

Glu Leu Lys Asn Ala Leu Arg
65                  70
```

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "A1;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 46
(D) OTHER INFORMATION: /note= "F1;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 51
(D) OTHER INFORMATION: /note= "V1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 52
(D) OTHER INFORMATION: /note= "S1;A1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 56
(D) OTHER INFORMATION: /note= "R1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 70
(D) OTHER INFORMATION: /note= "K1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 72
(D) OTHER INFORMATION: /note= "G1;E1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 73
(D) OTHER INFORMATION: /note= "D1;G1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 74
(D) OTHER INFORMATION: /note= "L1;S1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

```
Pro Asn Ser Glu Asn Gly Ala Cys Tyr Tyr Gly Asp Phe Ile Asp Tyr
1               5                   10                  15

Glu Glu Leu Arg Glu Gln Leu Ser Ser Val Ser Ser Phe Glu Arg Phe

            20                  25                  30

Glu Ile Phe Pro Lys Glu Ser Ser Trp Pro Asn His Thr Phe Asn Gly
        35                  40                  45

Val Thr Val Ser Cys Ser His Arg Gly Lys Ser Ser Phe Tyr Arg Asn
    50                  55                  60

Leu Leu Trp Leu Thr Lys Lys Gly Asp Leu Tyr Pro Lys Leu
65                  70                  75
```

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "E6;K3;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "S6;P3;T1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "E7;L1;Q1;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "T5.5;I2.5;S1;P1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "N9;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "N9;K1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "N9;Y1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "-8;I1;R1"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "T9;Q1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= "R8;K1;S1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "K9;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "S5.5;G2.5;I1;R1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "I9;V1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "I6;L3;M1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "P9;W1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "R6.5;Q3.5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "F6.5;W3.5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "Y8;V1;H1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "A6.5;T3.5"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 30
(D) OTHER INFORMATION: /note= "T9;I1"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "D3;K2.5;Q2.5;E1;G1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

```
Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Xaa Thr Arg
1               5                   10                  15

Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Ala Thr Gly Asp
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /note= "T60;K26;D7;A7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= "I93;N7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= "T67;K20;A13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5

(D) OTHER INFORMATION: /note= "Q46;E40;T7;K7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= "L86;T7;I7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 8
(D) OTHER INFORMATION: /note= "S53;D33;N7;G7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 9
(D) OTHER INFORMATION: /note= "G40;K26;V20;A7;S7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 10
(D) OTHER INFORMATION: /note= "L87;I13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 11
(D) OTHER INFORMATION: /note= "K85;N15"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 12
(D) OTHER INFORMATION: /note= "N60;G20;D20"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /note= "S67;L33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 14
(D) OTHER INFORMATION: /note= "E86;G14"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 15
(D) OTHER INFORMATION: /note= "E66;G27;K7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "K93;S7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 18
(D) OTHER INFORMATION: /note= "E67;T13;A13;K7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 19
(D) OTHER INFORMATION: /note= "K67;E33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 21
(D) OTHER INFORMATION: /note= "E40;A34;D13;K13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "K40;A26;E13;N7;D7;T7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 23
(D) OTHER INFORMATION: /note= "A87;V13"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "K66;D13;E7;N7;Q7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 26
(D) OTHER INFORMATION: /note= "C93;A7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 27
(D) OTHER INFORMATION: /note= "S93;N7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 28
(D) OTHER INFORMATION: /note= "E60;D12;K7;Q7;A7;T7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "E40;K20;A13;D13;T7;S7"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "T93;S7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "N60;K26;T20;D7;A7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 35
(D) OTHER INFORMATION: /note= "K72;E7;T7;Q7;S7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 36
(D) OTHER INFORMATION: /note= "S40;D33;G13;E7;N7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 37
(D) OTHER INFORMATION: /note= "S47;E20;G13;N13;K7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 38
(D) OTHER INFORMATION: /note= "H93;Q7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 39
(D) OTHER INFORMATION: /note= "A80;T20"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 40
(D) OTHER INFORMATION: /note= "E46;D27;Q20;V7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "I40;K33;V13;A7;L7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 44
(D) OTHER INFORMATION: /note= "Q40;E26;A20;V7;D7"

(ix) FEATURE:
(A) NAME/KEY: Modified-site

(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "G40;N34;D13;S13"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
Thr Leu Ile Thr Gln Lys Leu Ser Gly Leu Lys Asn Ser Glu Glu Leu
1               5                   10                  15

Lys Glu Lys Ile Glu Lys Ala Lys Lys Cys Ser Glu Glu Phe Thr Asn
            20                  25                  30

Lys Leu Lys Ser Ser His Ala Glu Leu Gly Ile Gln Gly
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= "N83;T17"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 17
(D) OTHER INFORMATION: /note= "A75;S25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 25
(D) OTHER INFORMATION: /note= "S83;A17"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 29
(D) OTHER INFORMATION: /note= "A92;S8"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 31
(D) OTHER INFORMATION: /note= "A50;T50"

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 34
(D) OTHER INFORMATION: /note= "A75;T25"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 41
(D) OTHER INFORMATION: /note= "V67;A33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 43
(D) OTHER INFORMATION: /note= "Q83;E17"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 50
(D) OTHER INFORMATION: /note= "A67;T33"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 56
(D) OTHER INFORMATION: /note= "S83;T17"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
Val His Val Gly Ala Asn Gln Asp Glu Ala Ile Ala Val Asn Ile Tyr
1               5                   10                  15

Ala Ala Asn Val Ala Asn Leu Phe Ser Gly Glu Gly Ala Gln Ala Ala
            20                  25                  30

Gln Ala Ala Pro Val Gln Glu Gly Val Gln Gln Glu Gly Ala Gln Gln
        35                  40                  45

Pro Ala Pro Ala Thr Ala Pro Ser Gln Gly Gly Val Asn
    50                  55                  60
```

We Claim:

1. A library of peptides comprising a known antigenic site, epitope or ligand on a protein molecule, wherein

(1) said library is optionally linked to a substantially invariant peptide domain or a core branched sequence;

(2) said library consists of an ensemble domain;

(3) the overall peptide length is about 8 to about 100 amino acids;

(4) the sequence of said ensemble domain is represented by a consensus formula;

(5) said ensemble is immunogenic, diagnostic for the epitope or is a therapeutic; and

(6) the consensus formula provides that each sequence position in the ensemble contain either a single amino acid or multiple amino acids, and that when a position contains multiple amino acids, (a) the identity and ratio of those amino acids being determined by the relative prevalence of amino acids in a consensus of known variant sequences for that epitope or ligand or (b) the identity of those amino acids being determined by the amino acids present in a consensus of known variant sequences for that epitope or ligand and the ratio of amino acids is equimolar.

2. The library of Claim 1, wherein said antigenic site, epitope or ligand is from a virus, bacterium, parasite, tumor antigen, allergen or other protein antigen, from a diagnostic marker site of an infectious agent or disease, from a therapeutically valuable ligand or from a helper T cell epitope or a cytotoxic T lymphocyte (CTL) epitope.

3. The library of Claim 1, wherein said antigenic site, epitope or ligand is from an envelope, core, NS1, NS3, NS4 or NS5 proteins of HCV; a gp120 V3 loop, gp41 envelope protein or the gp40 envelope protein of HIV; an envelope protein of HTLV I/II; an HA protein or a mutant HA domain of influenza A virus; an major outer membrane protein of Chlamydia

trachomtis; a neuraminidase, SAPA or CRA site of T. cruzi; an envelope B cell site or helper T cell site of type 2 Dengue virus; an M protein of streptococcus; a carbohydrate recognition site of selectin; a promiscuous helper T cell sites; a CTL epitope from HIV gag specific for HLA-B27; or an OspA, OspB, OspC or flagellin protein of Lyme disease.

4. The library of Claim 1 wherein said ensemble domain is represented by an SSAL library sequence provided in any one of Figs. 2-11, 13-34 (SEQ ID NO:1-43).

5. A peptide composition comprising the library of Claim 1.

6. The peptide composition of Claim 5 which is immunogenic.

7. A method of detecting antibodies associated with an infectious agent or a disease state which comprises using an effective amount of a peptide composition according to Claim 5 in an immunoassay procedure.

8. A method of detecting infection or a disease state which comprises contacting an effective amount of a peptide composition of Claim 5 with a body fluid, tissue or tissue extract in an immunoassay procedure for a time sufficient to form a complex between said peptide composition and any antibody in said fluid, said tissue, or said tissue extract, and subjecting said complex to a detecting means.

9. The method of Claim 7 wherein said immunoassay procedure is an ELISA or a PHA procedure.

10. The method of Claim 8 wherein said immunoassay procedure is an ELISA or a PHA procedure.

11. A kit for detection or diagnosis of an infectious agent or a disease state comprising a first container adapted to contain the peptide composition of Claim 5.

12. The kit of Claim 11 wherein said kit is an ELISA or PHA test kit.

13. A method of treating a disease which comprises administering an effective amount of a composition of Claim 5 to a patient for a time sufficient to elicit an efficacious result.

14. A pharmaceutical composition comprising the library of Claim 1, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier.

15. A method of vaccination which comprises administering an amount of the library of Claim 1 to a mammal effective to produce a protective immune response against an infectious agent which provided the basis of the library.

16. A vaccine composition comprising a library of Claim 1, or a salt thereof, and immunologically acceptable carrier.

17. A method of making an open diagnostic, vaccine or therapeutic composition for a mutatable infectious agent or for a protein site of known diversity which comprises

(1) aligning a collection of primary amino acid sequences for a related family of antigens, epitopes, diagnostic markers or therapeutic sites from said mutatable infectious agent or said protein site;

(2) identifying invariant and variant amino acid positions in said alignment;

(3) calculating a consensus formula for the sequence of a structured synthetic antigen library (SSAL);

(4) preparing said SSAL from said consensus formula.

18. The method of Claim 17 wherein said consensus formula provides that each sequence position in the alignment contains either a single amino acid or multiple amino acids, and that when a position contains multiple amino acids, (a) the identity and ratio of said multiple amino acids being determined by the relative prevalence of amino acids in said alignment, (b) the identity of said multiple amino acids being determined by the amino acids present in said alignment and the ratio of said multiple amino acids is equimolar, or (c)

the identity of said multiple amino acids being determined by the amino acids present in said alignment and the ratio of said multiple amino acids is weighted towards the predominant amino acid at said position.

19. The method of Claim 17 wherein said mutatable infectious agent is from a virus, bacterium or parasite.

20. The method of Claim 17 wherein said protein site is from a tumor antigen, allergen or other protein antigen, from a helper T cell epitope or a cytotoxic T lymphocyte (CTL) epitope.

21. The method of Claim 17, wherein said family of antigens, epitopes, diagnostic markers or therapeutic sites is from an envelope, core, NS1, NS3, NS4 or NS5 proteins of HCV; a gp120 V3 loop, gp41 envelope protein or the gp40 envelope protein of HIV; an envelope protein of HTLV I/II; an HA protein or a mutant HA domain of influenza A virus; an major outer membrane protein of Chlamydia trachomtis; a neuraminidase, SAPA or CRA site of T. cruzi; an envelope B cell site or helper T cell site of type 2 Dengue virus; an M protein of streptococcus; a carbohydrate recognition site of selectin; a promiscuous helper T cell sites; a CTL epitope from HIV gag specific for HLA-B27; or an OspA, OspB, OspC or flagellin protein of Lyme disease.

$$[AA_{1j}][AA_{2j}][AA_{3j}][AA_{4j}][AA_{5j}] \cdots\cdots\cdots\cdots [AA_{ij}]$$

$$\sum_{j=1}^{n} AA_{ij} = 1,$$

Figure 1

	S_{8}	G_{8}	K_{6}	P_{8}	A_{8}	I_{6}	I_{6}	P_{10}	D_{10}	R_{8}	E_{10}	V_{9}	L_{10}	Y_{10}	R_{3}	E_{8}	F_{10}	D_{10}	E_{10}	M_{10}	E_{10}	E_{10}	C_{10}	S_{4}	Q_{4}	H_{8}	L_{8}	P_{8}	Y_{8}	I_{9}	E_{10}	Q_{8}	G_{10}	M_{8}	M_{4}	L_{8}	A_{10}	E_{10}	Q_{8}	F_{8}	K_{10}	Q_{8}	K_{10}	A_{8}	L_{8}	G_{10}	L_{10}
	N_{2}	D_{1}	R_{4}	V_{1}	V_{2}	V_{4}	V_{2}			K_{2}		I_{1}			Q_{5}	A_{2}								A_{6}	S_{6}	K_{1}	A_{2}	A_{2}	L_{2}	F_{1}		E_{2}		Q_{2}	Q_{4}	M_{1}			M_{2}	L_{2}		S_{2}		I_{2}	Q_{2}		
		Q_{1}		A_{1}			A_{2}								E_{2}											R_{1}									R_{2}	I_{1}											
M62321	S	G	K	P	A	I	I	P	D	R	E	V	L	Y	R	E	F	D	E	M	E	E	C	S	Q	H	L	P	Y	I	E	Q	G	M	M	L	A	E	Q	F	K	Q	K	A	L	G	L
M67463	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
M58335	.	.	R	.	.	.	V	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	A	S	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.
D90208	.	.	R	.	.	V	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	A	S	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.
M84754	.	.	.	.	.	V	V	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	A	S	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.
D01221	N	D	R	V	V	V	A	.	.	K	.	I	.	.	E	A	.	.	.	.	.	.	.	A	S	K	A	A	L	.	.	E	.	Q	R	M	.	.	M	L	.	S	.	I	Q	.	.
D00944	N	Q	R	A	V	V	A	.	.	K	.	.	.	.	E	A	.	.	.	.	.	.	.	A	S	R	A	A	L	.	.	E	.	Q	R	I	.	.	M	L	.	S	.	I	Q	.	.
JANG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	A	S	.	.	.	.	F	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.

Figure 2

$S_{10} T_{10} N_{10} P_{9} K_{10} P_{10} Q_{10} K_{3} K_{10} N_{3} K_{10} R_{10} N_{10} T_{10} N_{10} R_{10} R_{10} P_{10} Q_{10} D_{10} V_{9} K_{10} F_{10} P_{10} G_{10} G_{10} G_{10} Q_{10} I_{10} V_{10}$

G_{1} (under P_9), R_{7} (under K_3), T_{7} (under N_3), I_{1} (under V_9)

M62321	S	T	N	P	K	P	Q	K	K	N	K	R	N	T	N	R	R	P	Q	D	V	K	F	P	G	G	G	Q	I	V
M67463	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M58335	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
D90208	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M84754	–	–	–	G	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
D01221	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
D00944	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
JANG	–	–	–	–	–	–	–	R	–	T	–	–	–	–	–	–	–	–	–	–	I	–	–	–	–	–	–	–	–	–

$G_{10} G_{10} V_{10} Y_{10} L_{10} L_{10} P_{10} R_{10} R_{10} G_{10} P_{10} R_{10} L_{10} G_{10} V_{10} R_{10} A_{10} T_{9} R_{10} K_{10} T_{10} S_{9} E_{10} R_{10} S_{10} Q_{10} P_{10} R_{10} G_{10} R_{10} R_{10}$

P_{1} (under T_9), W_{1} (under S_9)

M62321	G	G	V	Y	L	L	P	R	R	G	P	R	L	G	V	R	A	T	R	K	T	S	E	R	S	Q	P	R	G	R	R
M67463	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M58335	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	P	–	–	–	–	–	–	–	–	–	–	–	–	–
D90208	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M84754	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	W	–	–	–	–	–	–	–	–	–
D01221	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
D00944	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
JANG	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Figure 3

	K_9	A_{10}	I_{10}	P_{10}	L_6	E_8	V_4	I_{10}	K_9	G_{10}	G_{10}	R_{10}	H_{10}	L_{10}	I_{10}	F_{10}	C_{10}	H_{10}	S_{10}	K_{10}	K_{10}	K_{10}	C_{10}	D_{10}	E_{10}	L_{10}	A_{10}	A_{10}	K_8	L_{10}	V_4	A_5	L_8	G_{10}	I_6	N_9	A_{10}	V_{10}	A_{10}	Y_{10}
	R_1				I_4	A_1	A_3		R_1																				A_2		S_3	G_5	M_2		L_3	H_1				
						S_1	T_1																								T_1				V_1					
							F_1																								R_2									
							Y_1																																	
M62321	K	A	I	P	L	E	V	I	K	G	G	R	H	L	I	F	C	H	S	K	K	K	C	D	E	L	A	A	K	L	V	A	L	G	I	N	A	V	A	Y
M67463	-	-	-	-	I	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M58335	-	-	-	-	I	-	A	-	R	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	G	-	-	-	-	-	-	-	-
D90208	-	-	-	-	I	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	G	-	-	L	-	-	-	-	-
M84754	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	-	H	-	-	-	-
D01221	-	-	-	-	-	A	F	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	R	G	M	-	V	-	-	-	-	-
D00944	R	-	-	-	-	S	Y	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	R	G	M	-	L	-	-	-	-	-
JANG	-	-	-	-	I	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	G	-	-	L	-	-	-	-	-

	Y_{10}	R_{10}	G_{10}	L_{10}	D_{10}	V_{10}	S_{10}	V_{10}	I_{10}	P_{10}	T_9	S_7	G_{10}	D_9	V_{10}	V_{10}	V_{10}	V_{10}	A_9	T_{10}	D_{10}	A_{10}	L_{10}	M_{10}	T_{10}	G_{10}	Y_5	T_{10}	G_{10}	D_{10}	F_{10}	D_{10}	S_{10}	V_{10}	I_{10}	D_{10}	C_{10}	N_{10}	T_8	C_8
											A_1	I_1		N_1					S_1								F_5												V_2	A_2
												Q_2																												
M62321	Y	R	G	L	D	V	S	V	I	P	T	S	G	D	V	V	V	V	A	T	D	A	L	M	T	G	Y	T	G	D	F	D	S	V	I	D	C	N	T	C
M67463	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	-	-
M58335	-	-	-	-	-	-	-	-	-	-	-	I	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D90208	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	-	-
M84754	-	-	-	-	-	-	-	-	-	-	A	-	-	N	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	-	-
D01221	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	A
D00944	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	V	A
JANG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	F	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 4

	A_{10}	R_9	P_{10}	D_{10}	Y_{10}	N_{10}	P_{10}	P_8	L_{10}	V_5	E_{10}	T_5	W_{10}	K_{10}	K_4	P_{10}	D_9	Y_{10}	E_5	P_{10}	P_9	V_8	V_{10}	H_8	G_{10}	C_{10}	P_8	L_{10}	P_{10}	P_9	P_5	K_8	S_4	P_8	P_{10}	V_5	P_{10}	P_{10}	P_{10}	R_{10}	K_4	K_7	R_{10}	T_9
		P_1						L_1		L_4		S_4			D_4		G_1		V_4		A_1	T_2		L_1			A_2			S_1	T_3	P_1	A_4	T_2		I_4					R_6	R_3		A_1
								V_1		I_1		P_1			R_2				Q_1					A_1							V_1	R_1	K_1											T_1
																															I_1		Q_1											
M62321	A	R	P	D	Y	N	P	P	L	V	E	T	W	K	K	P	D	Y	E	P	P	V	V	H	G	C	P	L	P	P	P	K	S	P	P	V	P	P	P	R	K	K	R	T
M67463	.	.	.	.	.	.	.	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.
M58335	.	.	.	.	.	.	.	.	.	L	.	S	.	.	D	.	.	.	V	.	.	.	.	.	.	.	.	.	.	.	I	.	A	.	.	I	.	.	.	.	R	.	.	.
D90208	.	.	.	.	.	.	.	.	.	L	.	S	.	.	D	.	.	.	V	.	.	.	.	.	.	.	.	.	.	S	T	.	A	.	.	I	.	.	.	.	R	.	.	.
M84754	.	.	.	.	.	.	.	.	.	L	.	P	.	.	D	.	.	.	V	.	.	.	.	.	.	.	.	.	.	.	V	.	A	.	.	I	.	.	.	.	R	.	.	.
D01221	.	.	.	.	.	.	.	V	.	I	.	.	.	.	R	.	G	.	.	.	.	T	.	L	.	.	A	.	.	.	T	P	Q	T	.	.	.	.	.	.	R	R	.	A
D00944	.	.	.	.	.	.	.	.	.	.	.	S	.	.	R	.	.	.	Q	.	A	T	.	A	.	.	A	.	.	.	.	.	K	T	.	T	.	.	.	.	R	R	.	.
JANG	.	P	.	.	.	.	.	.	.	L	.	S	.	.	D	.	.	.	V	.	.	.	.	H	.	.	.	.	.	.	T	.	A	.	.	I	.	.	.	.	R	R	.	.

Figure 5

$C_{10}E_8\ P_{10}E_{10}P_{10}D_{10}V_8\ A_8\ V_{10}L_{10}T_9\ S_{10}M_{10}L_{10}T_{10}D_{10}P_{10}S_{10}H_{10}\ I_{10}T_{10}A_{10}E_{10}A_6\ A_{10}G_5\ R_{10}R_{10}L_{10}A_{10}R_{10}G_{10}S_{10}P_{10}P_{10}S_{10}$

		D_2					T_2	E_1			A_1													T_4		K_3										
								D_1																		A_2										
M62321	C	E	P	E	P	D	V	A	V	L	T	S	M	L	T	D	P	S	H	I	T	A	E	A	A	G	R	R	L	A	R	G	S	P	P	S
M67463	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M58335	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	K	-	-	-	-	-	-	-	-	-	-
D90208	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	K	-	-	-	-	-	-	-	-	-	-
M84754	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	K	-	-	-	-	-	-	-	-	-	-
D01221	-	D	-	-	-	-	T	E	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-
D00944	-	D	-	-	-	-	T	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	A	-	-	-	-	-	-	-	-	-	-

Figure 6

	C_{16}	L_{12} I_2 V_2	T_{16}	V_7 T_4 I_5	P_{16}	A_{12} V_4	S_{16}	A_{16}	Y_{12} V_2 A_2	Q_5 E_{11}	V_{16}	R_{11} H_3 K_2	N_{16}	S_5 V_7 I_4	T_3 S_{13}	G_{12} S_2 T_2	L_5 I_7 S_2 G_2	Y_{16}	H_{12} Y_2 M_2	V_{14} A_2	T_{16}	N_{16}	D_{16}	C_{16}	P_5 S_9 T_2	N_{16}	S_{10} A_2 N_2 D_2
M62321	C	L	T	V	P	A	S	A	Y	Q	V	R	N	S	T	G	L	Y	H	V	T	N	D	C	P	N	S
M67463	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	-	-	-	-	-	-	-	-	-
M58335	-	-	-	T	-	-	-	-	-	E	-	H	-	V	S	-	I	-	-	-	-	-	-	-	S	-	A
D90208	-	-	-	I	-	-	-	-	-	E	-	-	-	V	S	-	I	-	-	-	-	-	-	-	S	-	-
M84754	-	-	-	I	-	-	-	-	-	E	-	H	-	V	S	-	I	-	-	-	-	-	-	-	S	-	-
D01221	-	V	-	-	-	V	-	-	V	E	-	-	-	I	S	S	S	-	Y	A	-	-	-	-	S	-	N
D00944	-	I	-	T	-	V	-	-	A	E	-	K	-	I	S	T	G	-	M	-	-	-	-	-	T	-	D
HC-J1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-	-	-	-	-	-	-	-	-	-
HC-J4	-	-	-	I	-	-	-	-	-	E	-	-	-	V	S	-	I	-	-	-	-	-	-	-	S	-	-
HCV-JH	-	-	-	I	-	-	-	-	-	E	-	-	-	V	S	-	I	-	-	-	-	-	-	-	S	-	-
HCV-J	-	-	-	I	-	-	-	-	-	E	-	-	-	V	S	-	I	-	-	-	-	-	-	-	S	-	-
HCV-BK	-	-	-	T	-	-	-	-	-	E	-	H	-	V	S	-	I	-	-	-	-	-	-	-	S	-	A
HCV-J6	-	I	-	T	-	V	-	-	A	E	-	K	-	I	S	T	G	-	M	-	-	-	-	-	T	-	D
HCV-J7	-	V	-	-	-	V	-	-	V	E	-	-	-	I	S	S	S	-	Y	A	-	-	-	-	S	-	N

	S_{16}	I_{15} V_1	V_{12} T_4	Y_{12} W_4	E_{12} Q_4	A_{12} L_4	A_{11} H_1 T_2 Q_2	D_{13} A_2 N_1	A_9 L_2 M_4 V_1	I_{12} V_4	L_9 M_7	H_{16}	T_{11} L_2 V_2 A_1	P_{16}	G_{16}	C_{16}	V_{16}	P_{16}	C_{16}	V_{12} E_4	R_{12} N_2 K_2	E_{12} D_2 V_2	G_9 S_2 N_4 D_1	N_{14} G_2	A_4 S_5 F_2 T_4 V_1	S_{14} L_2	R_{15} H_1	C_{16}
M62321	S	I	V	Y	E	A	A	D	A	I	L	H	T	P	G	C	V	P	C	V	R	E	G	N	A	S	R	C
M67463	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M58335	-	-	-	-	-	-	-	-	L	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-
D90208	-	-	-	-	-	-	-	-	M	-	M	-	-	-	-	-	-	-	-	-	-	-	S	-	F	-	-	-
M84754	-	-	-	-	-	-	-	-	M	-	M	-	-	-	-	-	-	-	-	-	-	-	N	-	S	-	-	-
D01221	-	-	T	W	Q	L	T	-	-	V	-	-	L	-	-	-	-	-	-	E	N	D	N	G	T	L	H	-
D00944	-	-	T	W	Q	L	Q	A	-	V	-	-	V	-	-	-	-	-	-	E	K	V	-	-	T	-	-	-
HC-J1	-	-	-	-	-	-	H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	V	-	-	-
HCV1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HC-J4	-	-	-	-	-	-	-	-	M	-	M	-	-	-	-	-	-	-	-	-	-	-	D	-	S	-	-	-
HCV-JH	-	-	-	-	-	-	-	-	V	-	M	-	A	-	-	-	-	-	-	-	-	-	N	-	S	-	-	-
HCV-J	-	-	-	-	-	-	-	-	M	-	M	-	-	-	-	-	-	-	-	-	-	-	S	-	F	-	-	-
HCV-BK	-	-	-	-	-	-	-	-	L	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	S	-	-	-
HCV-J6	-	-	T	W	Q	L	Q	A	-	V	-	-	V	-	-	-	-	-	-	E	K	V	-	-	T	-	-	-
HCV-J7	-	-	T	W	Q	L	T	N	-	V	-	-	L	-	-	-	-	-	-	E	N	D	N	G	T	L	-	-

Figure 7

	C_{16}	V_{12}	R_{12}	E_{12}	G_9	N_{14}	A_4	S_{14}	R_{15}	C_{16}	W_{16}	V_{12}	A_{12}	M_3	T_{14}	P_{16}	T_{12}	V_9	A_{16}	T_5	R_{12}	D_5	G_5	K_5	L_5	P_{12}	A_3	T_{11}	Q_5	L_{10}	R_{16}	R_{12}	H_{16}	I_7	D_{16}	L_{12}	L_{12}	V_{16}	G_{12}	S_7	A_{16}	T_9	L_9	C_{16}
		E_4	K_2	V_2	S_2	G_2	S_5	L_2	H_1			I_4	Q_2	V_6	S_2		N_4	L_7		A_7	Q_2	N_7	A_2	S_5	V_3	L_4	T_{13}	Q_2	T_7	I_6		T_4		V_9		M_4	I_2		M_4	A_8		A_7	F_5	
			N_2	D_2	N_4		F_2						P_2	L_7						V_4	K_2	Q_2	S_2	T_2	I_4			H_1	G_2								V_2			T_1			V_1	
					D_1		T_4															H_2	V_2	G_4	A_4			A_1	N_1														A_1	
							V_1																P2					R_1	S_1															
																							R_2																					
																							N_1																					
M62321	C	V	R	E	G	N	A	S	R	C	W	V	A	M	T	P	T	V	A	T	R	D	G	K	L	P	A	T	Q	L	R	R	H	I	D	L	L	V	G	S	A	T	L	C
M67463	–	–	–	–	–	–	–	–	–	–	–	–	–	V	–	–	–	–	–	–	–	–	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M58335	–	–	–	–	–	–	S	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	V	T	I	–	T	–	T	I	–	–	–	V	–	–	–	–	–	A	–	A	F	–
D90208	–	–	–	–	S	–	F	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	S	S	I	–	T	–	T	I	–	–	–	V	–	–	–	–	–	A	–	A	–	–
M84754	–	–	–	–	N	–	S	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	N	S	V	–	T	A	T	I	–	–	–	V	–	–	–	–	–	A	–	A	F	–
D01221	–	E	N	D	N	G	T	L	H	–	–	I	Q	V	–	–	N	–	–	V	K	H	R	G	A	L	T	R	S	–	–	T	–	V	–	M	I	–	M	A	–	–	A	–
D00944	–	E	K	V	–	–	T	–	–	–	–	I	P	V	S	–	N	–	–	V	Q	Q	P	G	A	L	T	Q	G	–	–	T	–	–	–	M	V	–	M	–	–	–	–	–
HC–J1	–	–	–	–	–	–	V	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
HCV1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
HCV–H	–	–	–	–	–	–	–	–	–	–	–	–	–	V	–	–	–	–	–	–	–	–	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
HC–J4	–	–	–	–	D	–	S	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	A	S	V	–	T	–	T	I	–	–	–	V	–	–	–	–	–	A	–	A	F	–
HCV–JH	–	–	–	–	N	–	S	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	A	S	V	–	T	–	T	–	–	–	–	V	–	–	–	–	–	T	–	A	F	–
HCV–J	–	–	–	–	S	–	F	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	S	S	I	–	T	–	T	I	–	–	–	V	–	–	–	–	–	A	–	A	–	–
HCV–BK	–	–	–	–	–	–	S	–	–	–	–	–	–	L	–	–	–	L	–	A	–	N	V	T	I	–	T	–	T	I	–	–	–	V	–	–	–	–	–	A	–	A	F	–
HCV–J8	–	E	K	V	–	–	T	–	–	–	–	I	P	V	S	–	N	–	–	V	Q	Q	P	G	A	L	T	Q	G	–	–	T	–	–	–	M	V	–	M	–	–	–	–	–
HCV–J7	–	E	N	D	N	G	T	L	–	–	–	I	Q	V	–	–	N	–	–	V	K	H	R	G	A	L	T	H	N	–	–	T	–	V	–	M	I	–	M	A	–	–	V	–

Figure 8

	F_{16}	T_{12}	F_{12}	S_{16}	P_{16}	R_{12}	R_{13}	H_{13}	W_{7}	T_{12}	T_{7}	Q_{16}	G_{3}	C_{16}	N_{16}	C_{16}	S_{16}	I_{16}	Y_{16}	P_{14}	G_{16}	H_{15}	I_{5}	T_{10}	G_{16}	H_{16}	R_{16}	M_{16}	A_{16}	W_{16}	D_{15}	M_{16}	M_{16}	M_{14}	N_{15}	W_{16}	S_{16}	P_{16}	T_{16}	T_{8}
		I_{3}	V_{3}			Q_{3}	H_{3}	Y_{3}	E_{5}	F_{4}	V_{7}		D_{11}							Q_{2}		T_{1}	L_{1}	S_{6}							N_{1}			L_{2}	S_{1}					A_{6}
		M_{1}	I_{1}			E_{1}			V_{2}		L_{2}		E_{2}										V_{6}																	L_{2}
									N_{2}																															
M62321	F	T	F	S	P	R	R	H	W	T	T	Q	G	O	N	C	S	I	Y	P	G	H	I	T	G	H	R	M	A	W	D	M	M	M	N	W	S	P	T	T
M67463	-	-	-	-	-	-	H	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	-	-	-	-	-	-	A
M58335	-	-	-	-	-	-	-	-	V	-	L	-	D	-	-	-	-	-	-	-	-	-	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D00208	-	-	-	-	-	-	-	Y	E	-	V	-	D	-	-	-	-	-	-	-	-	-	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M84754	-	-	-	-	-	-	-	Y	E	-	V	-	D	-	-	-	-	-	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
D01221	-	M	V	-	-	Q	-	-	N	F	-	-	E	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	L	S	-	-	-	-	L
D00944	-	I	V	-	-	Q	H	-	-	F	V	-	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
HC-J1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
HCV1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-H	-	-	-	-	-	-	-	-	-	-	-	-	D	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
HC-J4	-	-	-	-	-	-	-	-	E	-	V	-	D	-	-	-	-	-	-	-	-	-	L	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-JH	-	-	-	-	-	-	-	-	E	-	V	-	D	-	-	-	-	-	-	-	-	-	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
HCV-J	-	-	-	-	-	-	-	Y	E	-	V	-	D	-	-	-	-	-	-	-	-	-	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-BK	-	-	-	-	-	-	-	-	V	-	L	-	D	-	-	-	-	-	-	-	-	-	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCV-J8	-	I	V	-	-	O	H	-	-	F	V	-	D	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
HCV-J7	-	I	I	-	-	E	-	-	N	F	-	-	E	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	L	-	-	-	-	-	L

Figure 9

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	V_{10}	D_{10}	A_5 G_5	E_3 H_2 D_2 S_1 T_1 Q_1	T_{10}	H_7 I_2 Y_1	V_7 S_1 T_1	T_6 S_3 V_1	G_{10}	G_9 Q_1	S_2 Q_1 R_2 A_2 N_1 T_1 E_1	A_6 Q_1 V_2 T_1	G_3 A_7	H_2 R_4 S_2 K_2	T_6 S_2 N_1 A_1	V_2 T_6 M_1 A_1	S_2 A_2 N_2 Q_2 H_1 R_1	G_4 S_3 R_2 T_1
M62321	V	D	A	E	T	H	V	T	G	G	S	A	G	H	T	V	S	G
M87463	–	–	–	–	–	–	–	–	–	–	N	–	–	R	–	T	A	–
M58335	–	–	G	D	–	–	–	–	–	–	A	Q	A	K	–	T	N	R
D90208	–	–	G	H	–	–	–	–	–	–	R	V	A	S	S	T	Q	S
M84754	–	–	G	S	–	I	–	S	–	–	T	V	A	R	–	T	H	S
D01221	–	–	–	T	–	Y	S	S	–	Q	E	–	–	R	–	–	A	–
D00944	–	–	–	Q	–	–	T	V	–	–	–	T	A	–	N	A	R	T
HC–J1	–	–	–	–	–	I	–	S	–	–	Q	–	A	R	A	M	–	–
HCV–J	–	–	G	H	–	–	–	–	–	–	R	V	A	S	S	T	Q	S
HCV–BK	–	–	G	D	–	–	–	–	–	–	A	Q	–	K	–	T	N	R

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
	F_2 L_8	V_7 A_2 T_1	S_7 G_3	L_5 M_3 W_2	L_4 F_6	A_3 S_3 T_4	P_3 Q_3 S_2 T_1 L_1	G_{10}	A_6 P_4	K_4 S_5 R_1	Q_{10}	N_4 K_6	V_1 I_8 L_1	Q_9 Y_1	L_{10}	I_8 V_2	N_{10}
M62321	F	V	S	L	L	A	P	G	A	K	Q	N	V	Q	L	I	N
M87463	L	–	G	–	–	T	–	–	–	–	–	–	I	–	–	–	–
M58335	L	–	–	M	F	–	S	–	P	S	–	K	I	–	–	–	–
D90208	L	–	–	W	L	S	Q	–	P	S	–	K	I	–	–	V	–
M84754	L	A	–	–	F	T	Q	–	–	S	–	K	I	–	–	–	–
D01221	–	A	G	–	F	T	T	–	–	–	–	–	L	Y	–	–	–
D00944	L	T	G	M	F	S	L	–	–	R	–	K	I	–	–	–	–
HC–J1	L	–	–	–	F	T	–	–	–	–	–	–	–	–	–	–	–
HCV–J	L	–	–	W	–	S	Q	–	P	S	–	K	–	–	–	V	–
HCV–BK	L	–	–	M	–	–	S	–	P	S	–	K	–	–	–	–	–

Figure 10

	Q_8 I_1 V_1	G_{10}	W_{10}	G_{10}	P_8 T_1 A_1	I_8 L_2	S_5 T_3 E_1 Q_1	Y_9 H_1	A_6 D_1 T_1 E_2	N_5 E_2 M_1 T_T D_1	• N_2	• V_2	G_5 S_1 P_1 A_1 T_2	S_6 E_1 D_1 N_2	G_5 R_1 S_1 I_1 D_1 P_1	P_4 L_1 S_2 Q_1 G_1 E_1	D_{10}	Q_7 E_1 M_2	R_{10}	P_{10}	Y_{10}	C_{10}
M62321	Q	G	W	G	P	I	S	Y	A	N	•	•	G	S	G	P	D	Q	R	P	Y	C
M87483	–	–	–	–	–	–	–	–	–	–	•	•	–	–	–	L	–	E	–	–	–	–
M58335	–	–	–	–	–	–	T	–	–	E	•	•	S	–	R	S	–	–	–	–	–	–
D90208	–	–	–	–	–	–	T	H	D	M	•	•	P	E	S	S	–	–	–	–	–	–
M84754	–	–	–	–	–	–	T	–	T	E	•	•	A	D	I	Q	–	–	–	–	–	–
D01221	I	–	–	–	T	L	E	–	E	T	N	V	T	N	D	G	–	M	–	–	–	–
D00944	V	–	–	–	A	L	Q	–	E	D	N	V	T	N	P	E	–	M	–	–	–	–

	W_{10}	H_{10}	Y_{10}	P_8 A_2	P_{10}	K_4 R_5 P_1	P_8 Q_2	C_{10}	G_9 T_1	I_9 V_1	V_{10}	P_9 S_1	A_{10}	K_5 S_4 R_1	S_6 E_1 Q_2 T_1	V_{10}	C_{10}	G_{10}	P_{10}	V_{10}	Y_{10}	C_{10}
M62321	W	H	Y	P	P	K	P	C	G	I	V	P	A	K	S	V	C	G	P	V	Y	C
M87483	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
M58335	–	–	–	–	–	P	Q	–	T	–	–	–	–	S	E	–	–	–	–	–	–	–
D90208	–	–	–	A	–	R	–	–	–	–	–	–	–	S	Q	–	–	–	–	–	–	–
M84754	–	–	–	A	–	R	–	–	–	–	–	–	–	S	Q	–	–	–	–	–	–	–
D01221	–	–	–	–	–	R	–	–	–	–	–	–	–	R	T	–	–	–	–	–	–	–
D00944	–	–	–	–	–	R	Q	–	–	V	–	S	–	S	–	–	–	–	–	–	–	–

Figure 11

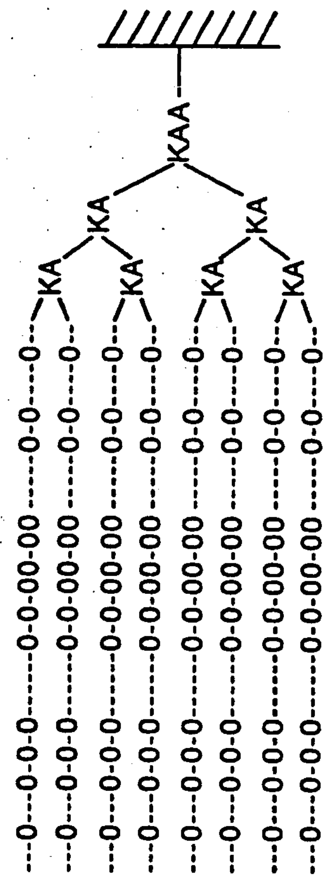


Figure 12 A

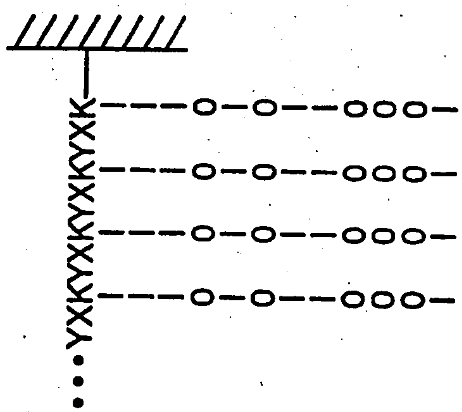


Figure 12 B

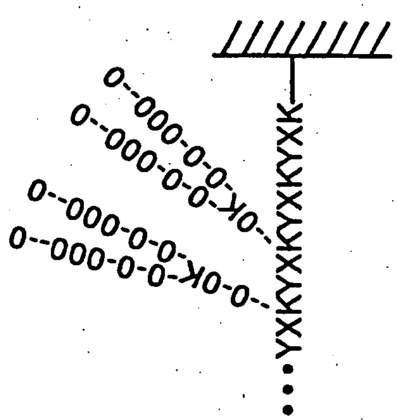


Figure 12 C

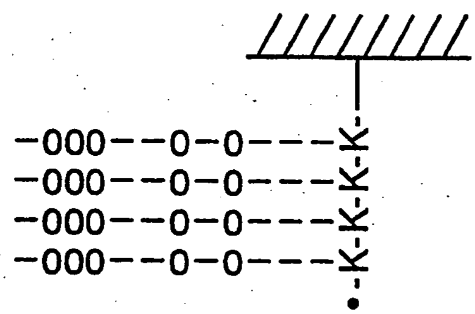


Figure 12 D

```
Con S        ESVEINCTRPNNNTRKSIHIGPGQAFYATGDM
Con A        KP-R--C----------V------T-------
Con B        -----------------------R---T--E-
Con D/S      ------------------PL------------
Con C        -----V------------R-----T-------
Con 1            -----------------R-------
Con 2            ------------R------------
Con 3            ------------P----R--------
Con 4            ----------R-T----RV--T--E-
Con 5            ----SK-V-RR------R---T-KQ-
Con 6            ----S----T--T-----V--R----
Con 7            ----Y----Q-T-------L-T-*K-
Con 8            ----Y-S--QRTS--Q---L-T-*M-
Con 9            ----YK---QRTP--L---L-T-*R-
Con 10           ----Y--K-Q-TP--L---L-T-*R-
V3mn         ---Q-------Y-K--R------R---T-KN-
V3sc         -A-----------TR-------R--------
ThaiB        -------------------L-----W-T--Q-
ThaiE        K---------S----T--P-----V--R----
IvCo         KP-R-----------EGVG-----T--K--N-
Rwanda       KA-K-------KT---GVR------W--R-N-
Haiti        -T-K---A-L-----R--PV---K-L-T----
V3WMJ        ----------Y--V-R-LS----R--*R-RE-
Uganda       ---T------YS---QGT-----R-YCTS-Y-
Brazil       ---V-----H---------V-W-RSLFT--E-
Uganda 88    ----------Y--IKQRTP--R---LFT-RR-
                          I                K
```

Figure 13 A

E_{12}	S_{11}	V_{16}	E_9	I_{16}	N_{15}	C_{16}	T_{15}	R_{16}	P_{14}	N_{12}	N_{13}	N_{15}	T_{13}	R_{14}	K_9	S_{11}	I_{10}	H_8	I_{12}	G_{16}	P_{14}	G_{16}	Q_9	A_{11}	F_{10}	Y_{12}	A_6	T_{14}	G_{13}	D_7
K_4	P_2		R_2		V_1		A_1		L_1	S_1	Y_1	T_1	K_1	K_1	R_3	R_2	V_3	P_4	L_2		W_1		R_6	T_3	W_2	F_2	T_7	R_1	K_1	E_3
	A_2		K_2						H_1	Y_3	K_1		V_1	T_1	Y_1	G_3	L_1	R_2	V_2		R_1		K_1	V_1	L_3	S_1	R_2	S_1	R_2	N_3
	T_1		Q_1								S_1		I_1		E_1		T_2	G_1						S_1	Y_1		K_1			Q_1
			T_1												Q_2			S_1												Y_1
			V_1																											R_1

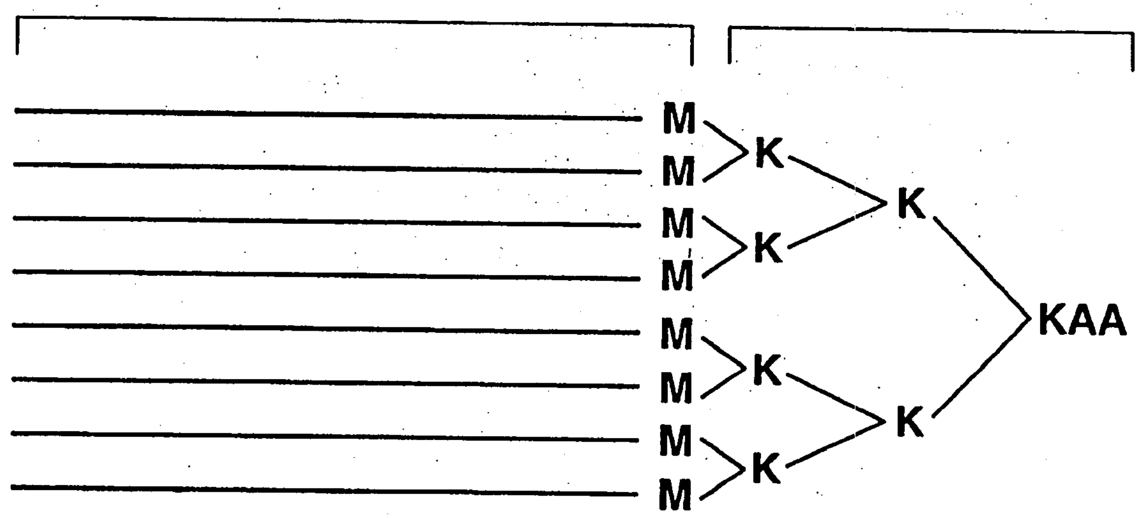


Figure 13 B

E_{6}	S_{6}	V_{10}	E_{7}	I_{10}	N_{10}	C_{10}	$T_{5.5}$	R_{10}	P_{10}	N_{9}	N_{9}	N_{9}	$-_{8}$	T_{9}	R_{8}	K_{9}	$S_{5.5}$	I_{9}	H_{3}	I_{6}	G_{10}	P_{9}	G_{10}	$R_{6.5}$	A_{10}	$F_{6.5}$	Y_{8}	$A_{6.5}$	T_{9}	G_{10}	D_{3}
K_{3}	P_{3}		L_{1}				$I_{2.5}$			S_{1}	K_{1}	Y_{1}	I_{1}	Q_{1}	K_{1}	R_{1}	$G_{2.5}$	V_{1}	T_{3}	L_{3}		W_{1}		$Q_{3.5}$		$W_{3.5}$	V_{1}	$T_{3.5}$	I_{1}		$K_{2.5}$
T_{1}	T_{1}		Q_{1}				S_{1}						R_{1}		S_{1}		I_{1}		P_{3}	M_{1}							H_{1}				$Q_{2.5}$
			V_{1}				P_{1}										R_{1}		R_{1}												E_{1}
																															G_{1}

Figure 13 C

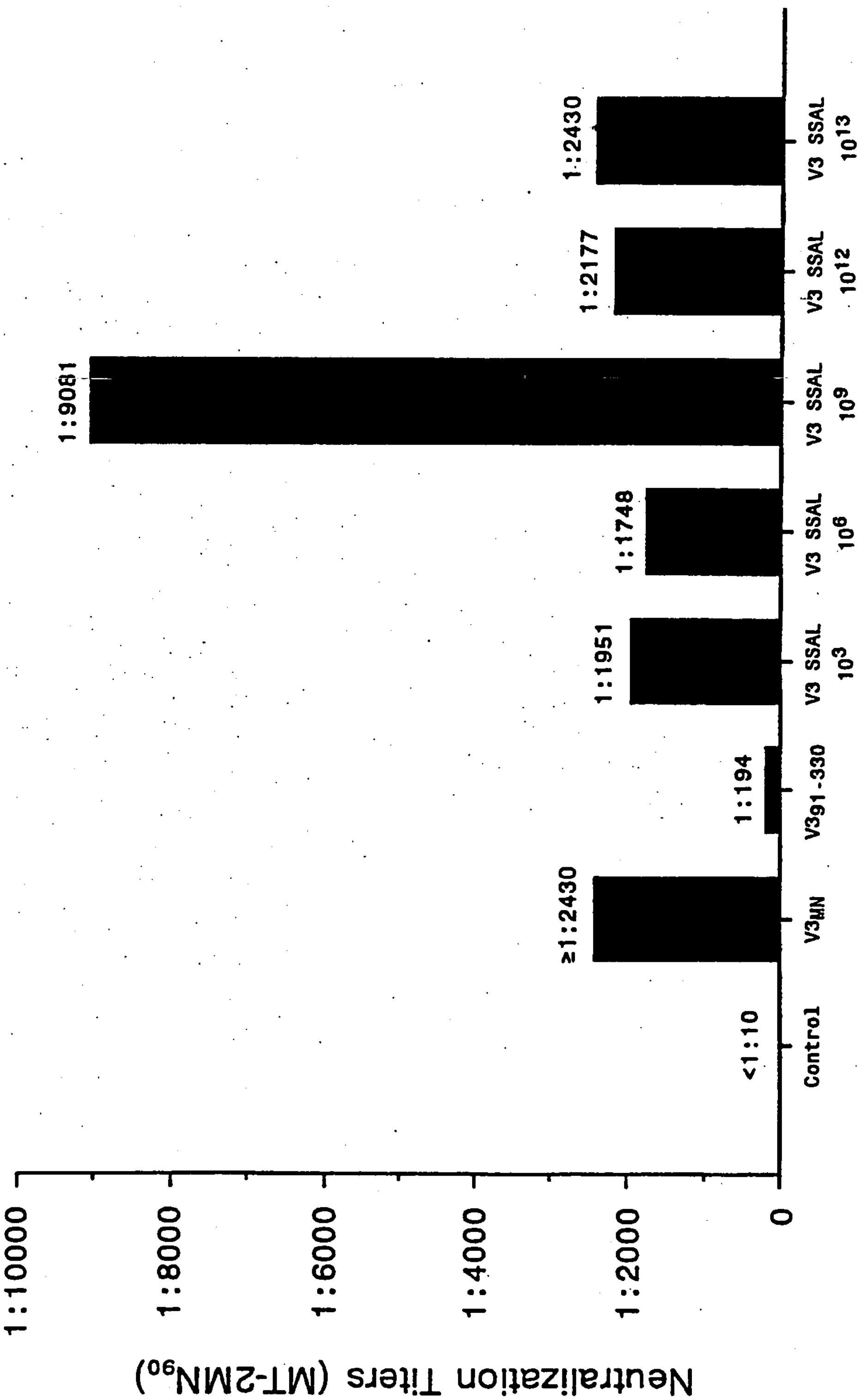

Neutralization Titers (MT-2MN90)
1:10000
1:8000
1:6000
1:4000
1:2000
0
<1:10
Control
≥1:2430
V3MN
1:194
V391-330
1:1951
V3 SSAL 10³
1:1748
V3 SSAL 10⁶
1:9081
V3 SSAL 10⁹
1:2177
V3 SSAL 10¹²
1:2430
V3 SSAL 10¹³

Figure 13 D

	R_{50}	I_{9}	L_{48}	A_{50}	V_{41}	E_{50}	R_{47}	Y_{47}	L_{49}	K_{27}	D_{47}	Q_{50}	Q_{49}	L_{48}	L_{48}	G_{48}	I_{42}	W_{49}	G_{49}	C_{49}	S_{48}	G_{50}	K_{49}	L_{43}	I_{47}	C_{50}	T_{48}	T_{50}	A_{22}	V_{50}	P_{48}	W_{50}	N_{50}	A_{26}	S_{48}
		V_{38}	Q_{1}		L_{5}		S_{1}	F_{1}	I_{1}	Q_{6}	E_{1}		R_{1}	R_{1}	M_{1}	N_{1}	F_{2}	L_{1}	W_{1}	R_{1}	K_{2}		R_{1}	H_{4}	V_{3}		Y_{3}		T_{20}		K_{2}			T_{8}	T_{2}
		L_{3}			I_{4}		T_{2}	L_{2}		G_{1}	N_{2}			I_{1}		S_{1}	M_{2}							I_{2}			P_{1}		N_{5}					S_{14}	
										R_{16}							G_{1}							A_{1}					F_{1}					R_{1}	
																	L_{3}												S_{2}					N_{1}	
HIVHXB2R	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	L	G	I	W	G	C	S	G	K	L	I	C	T	T	A	V	P	W	N	A	S
HIVSF1703	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	N	.	.	.	.	S	.
HIVU455	.	V	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	S	.
HIVZ321	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	I	.	.	P	.	N	.	.	.	.	S	.
HIVD687	.	V	.	.	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	S	.
HIVJRCSF	.	V	.	.	.	.	.	.	.	.	.	.	.	.	M	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.
HIVJRFL	.	V	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVALA1	.	V	.	.	.	.	.	.	.	R	.	.	.	.	.	E	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
HIVBRVA	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVJH3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVSC	.	V	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.
HIVBAL1	.	V	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVYU2	.	V	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	T	.
HIVMN	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	F	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
HIVLAI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVNL43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVMFA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVCAM1	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVNY5CG	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
HIVADA	.	V	.	.	L	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVJFL	.	V	.	.	.	.	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.
HIVSIMI84	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	T	.	.	.	.	.	.
HIVD31	.	V	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIVSF162	.	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 14 A

HIVHXB2R	R I L A V E R Y L K D Q Q L L G I W G C S G K L I C T T A V P W N A S
HIVOYI	. V . . . . T . . . . . .
HIVSF33	. V . . . . R . . . . T . . . . T .
HIVCDC4	. . . . F . . . .
HIVSF2	. V . . . . R . . . .
HIVSF2B13	. V . . . . R . . . . T .
HIVHAN	. V . . . . R . . . . T . . . . . .
HIVRF	. V . . . . R . . . . T . . . . . .
HIVWMJ2	. V . . . . R . . . . T . . . . . .
HIVTB132	. V . . . . T . . . . . .
HIVGUN	. V . . I . . F . R . . . . T . . . . T .
HIVRJS	. V . . . . T .
HIVSBA	. V . . . . R . . . .
HIVSBC	. V . . . . R . . . . S .
HIVSBB	. V . . L . . . .
HIVD747	. V . . I . . . . . E . . . . T . . . . S .
HIVD760	. V . . I . . . . M . . . . S .
HIVD757	. V . . I . . . . R . . . . N . . . . S .
HIVJY1	. V . . . . S . . . . H . . . . T . . . . S .
HIVNDK	. V . . . . R . . . . R H . . . . N . . . . S .
HIVMAL	. V . . . . Q . . R . . . M . . . . . . H . . . . F . . . . S .
HIVELI	. . . . H . . . . N . . . . S .
HIVZ2Z6	. . . . T . . . . S .
HIVTN243	. V . . . . G L W . . . . I . . . . S T
HIVMVP5180	. L Q . L . T L I Q N . . R . N L . . . K . . . . . Y . S . K . . T .
HIVANT70	. L . . L . T L . Q N . . . . S L . . . K . . . V . Y . S . K . . R T
CPZGAB	. L . . . . Q . . . I . . L . . . . . . A V . Y . T . . . . N .

Figure 14 B

	R_{20}	V_{20}	T_{19}	A_{20}	I_{20}	E_{20}	K_{20}	Y_{20}	L_{20}	Q_{1}	D_{19}	Q_{20}	A_{20}	R_{2}	L_{19}	N_{20}	S_{14}	W_{20}	G_{20}	C_{20}	A_{20}	F_{20}	R_{20}	Q_{20}	V_{20}	C_{20}	H_{20}	T_{20}	T_{18}	V_{20}	P_{20}	W_{19}	V_{7}	N_{20}	D_{10}	S_{11}
			S_{1}							K_{18}	H_{1}			Q_{16}	x_{1}		A_{6}												S_{1}			R_{1}	A_{1}		E_{4}	T_{8}
										A_{1}				L_{1}															x_{1}				P_{12}		A_{6}	N_{1}
														x_{1}																						
HIV2ROD	R	V	T	A	I	E	K	Y	L	Q	D	Q	A	R	L	N	S	W	G	C	A	F	R	Q	V	C	H	T	T	V	P	W	V	N	D	S
HIV2NIHZ	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.	.	.	.	.	.	T
HIV2ISYR	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
HIV2ST	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
HIV2BEN	.	.	.	.	.	.	.	.	.	K	H	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIV2CAM2	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	E	.
HIV2D194	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIV2GH1	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HIV2D205	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	E	T
HIV2UC1	.	.	.	.	.	.	.	.	.	K	.	.	.	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	E	T
SIVMM251	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	.
SIVMM32H	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	.
SIVMM1A11	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	.
SIVMM142	.	.	S	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	.
SIVMM239	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	.
SIVMNE	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	A	N
SIVSMMH4	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	E	T
SIVSMMH9	.	.	.	.	.	.	.	.	.	K	.	.	.	x	x	.	.	.	.	.	.	.	.	.	.	.	.	.	x	.	.	.	P	.	.	T
SIVSMMPBJ	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	R	P	.	.	T
SIVSTM	.	.	.	.	.	.	.	.	.	K	.	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	P	.	.	.

Figure 15

16 A (gp46) $L_1N_1T1E_2P_2S_1Q_2L1P_2P_2T_2A_1P_2P_2L_2L_1P_1H_1S_2W_1L_2D_1H_2I_1L_2E_1P_2S_2I_1P_1W_2K_1S_1K_2L_1L_2T_1L_1V_1Q_2L_2T_2L_2Q_2S_2$

$I_1T_1S_1$ T_1 P_1 S_1 $V_1H_1D_1$ D_1 E_1 V1 T_1 T_1S_1 T_1T_1 I_1 $K_1F_1I_1$

HTLV-1 L N T E P S Q L P P T A P P L L P H S W L D H I L E P S I P W K S K L L T L V Q L T L Q S

HTLV-2 I T S - - T - P - - - S - - - V H D - D - E - V - T - - T S - T T - I - K F I - - - - - - -

16 B (gp46) $C_2F_1D_1P_2Q_1I_1Q_2A_2I_2V_1S_1S_1P_1C_2H_2N_2S_2L_1I_2L_2P_2P_2S_2L_2S_1P_2V_2P_2T_1L_2G_1S_1R_2S_1R_2R_2A_2$

Y1Q1 R1L1 T1T1D1N1 I1 A1 P1 A1T1 R1

HTLV-1 C F D P Q I Q A I V S S P C H N S L I L P P S L S P V P T L G S R S R R A

HTLV-2 - Y Q - R L - - - T T D N - - - - I - - - - - - A - - - P - A T - R - - -

16 C (gp21) $V_2D_2K_2D_2I_2S_2Q_2L_2T_2Q_2A_2I_2V_2K_2H_2H_2K_1H_2L_1L_2K_1I_1A_2N_1Y_2A_2A_2Q_2H_2R_2R_2G_2L_2D_2L_2L2F_2W_2E_2Q_2G_2G_2L_2C_2K_2A_2L_1Q_2E_2Q_2C_2R_1F_2P_1H_2I_2T_1H_2S_1H_2V_2P_1I_1L_2Q_2E_2R_2P_2P_2L_2E_2$

H_1 Q_1 I_1 R_1V_1 Q_1 I_1 C_1 L1 S_1 T_1 S_1V_1

HTLV-1 V D K D I S Q L T Q A I V K H H K H L L K I A N Y A A Q H R R G L D L L F W E Q G G L C K A L Q E Q C R F P H I T H S H V P I L Q E R P P L E

HTLV-2 - - - - - - H - - - - - - - - - - Q - I - R V - Q - I - - - - C - L - - S - T - - S V - - - - - - - - -

Figure 16

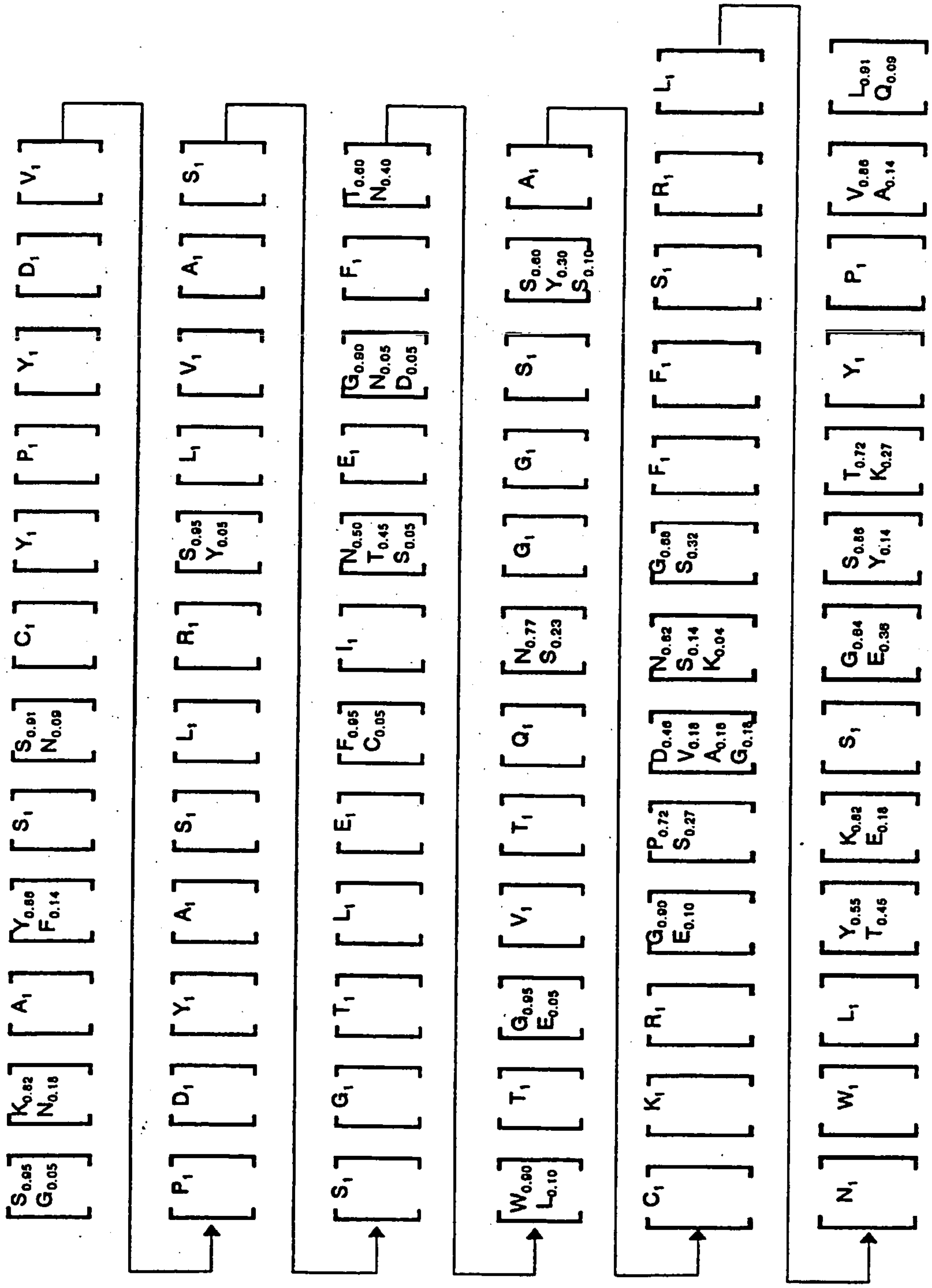

S0.95 G0.05
K0.82 N0.18
A1
Y0.86 F0.14
S1
S0.91 N0.09
C1
Y1
P1
Y1
D1
V1
P1
D1
Y1
A1
S1
L1
R1
S0.95 Y0.05
L1
V1
A1
S1
S1
G1
T1
L1
E1
F0.95 C0.05
I1
N0.50 T0.45 S0.05
E1
G0.90 N0.05 D0.05
F1
T0.60 N0.40
W0.90 L0.10
T1
G0.95 E0.05
V1
T1
Q1
N0.77 S0.23
G1
G1
S1
S0.60 Y0.30 S0.10
A1
C1
K1
R1
G0.90 E0.10
P0.72 S0.27
D0.46 V0.18 A0.18 G0.18
N0.82 S0.14 K0.04
G0.68 S0.32
F1
F1
S1
R1
L1
N1
W1
L1
Y0.55 T0.45
K0.82 E0.18
S1
G0.64 E0.36
S0.86 Y0.14
T0.72 K0.27
Y1
P1
V0.86 A0.14
L0.91 Q0.09

Figure 17

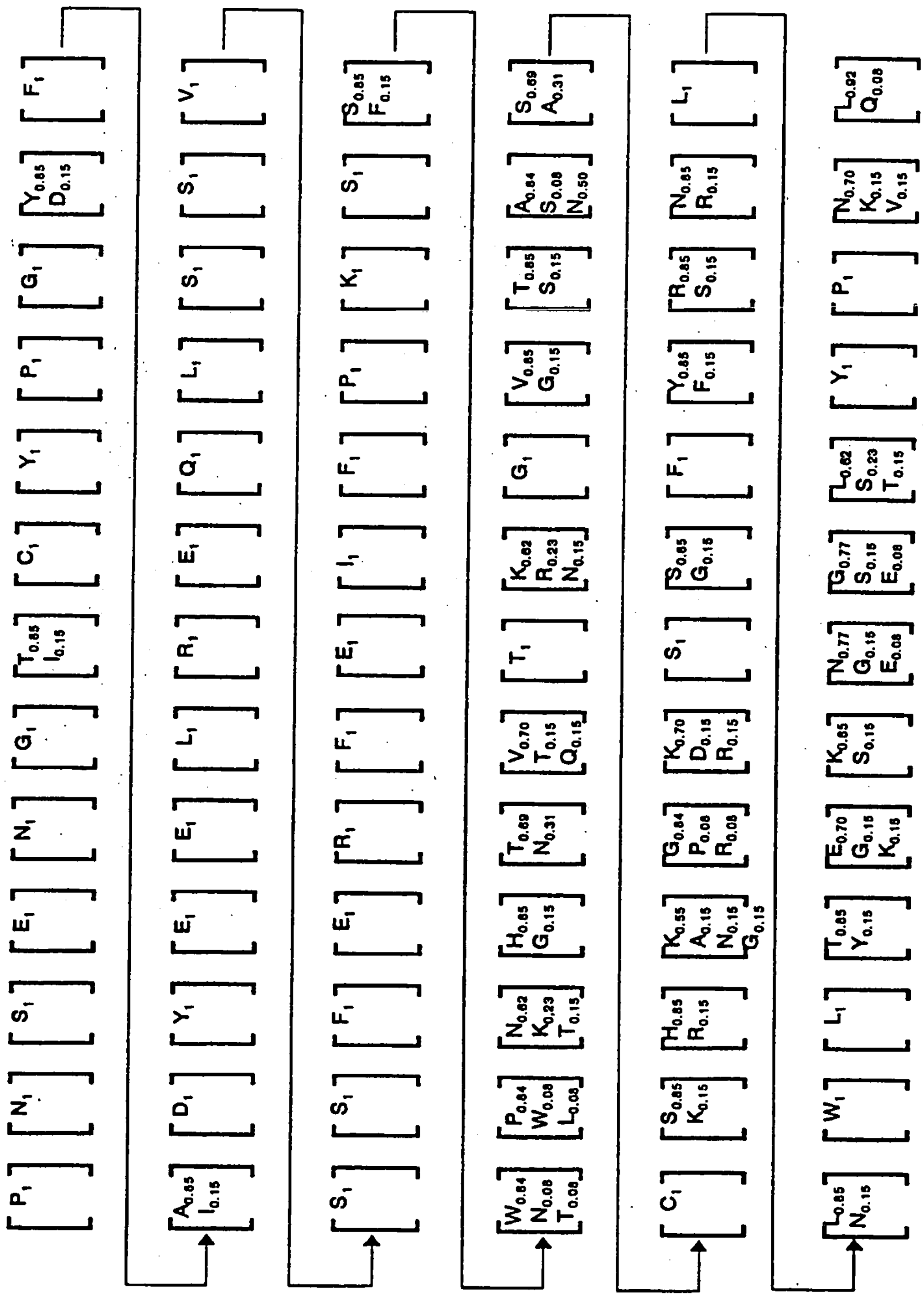


Figure 18 A

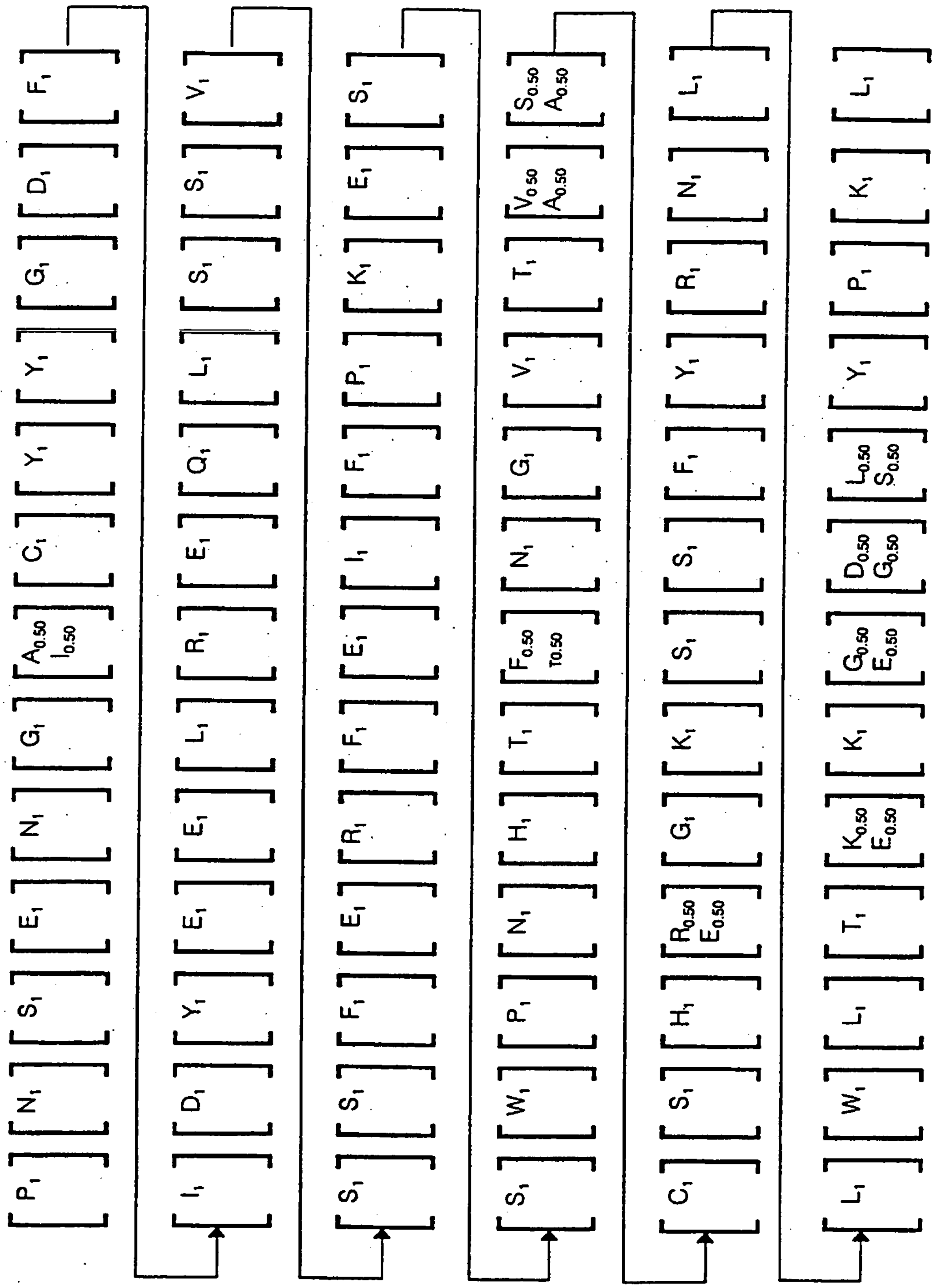

P_1 N_1 S_1 E_1 N_1 G_1 $A_{0.50}$ $I_{0.50}$ C_1 Y_1 Y_1 G_1 D_1 F_1
I_1 D_1 Y_1 E_1 E_1 L_1 R_1 E_1 Q_1 L_1 S_1 S_1 V_1
S_1 S_1 F_1 E_1 R_1 F_1 E_1 I_1 F_1 P_1 K_1 E_1 S_1
S_1 W_1 P_1 N_1 H_1 T_1 $F_{0.50}$ $T_{0.50}$ N_1 G_1 V_1 T_1 $V_{0.50}$ $A_{0.50}$ $S_{0.50}$ $A_{0.50}$
C_1 S_1 H_1 $R_{0.50}$ $E_{0.50}$ G_1 K_1 S_1 S_1 F_1 Y_1 R_1 N_1 L_1
L_1 W_1 L_1 T_1 $K_{0.50}$ $E_{0.50}$ K_1 $G_{0.50}$ $E_{0.50}$ $D_{0.50}$ $G_{0.50}$ $L_{0.50}$ $S_{0.50}$ Y_1 P_1 K_1 L_1

Figure 18 B

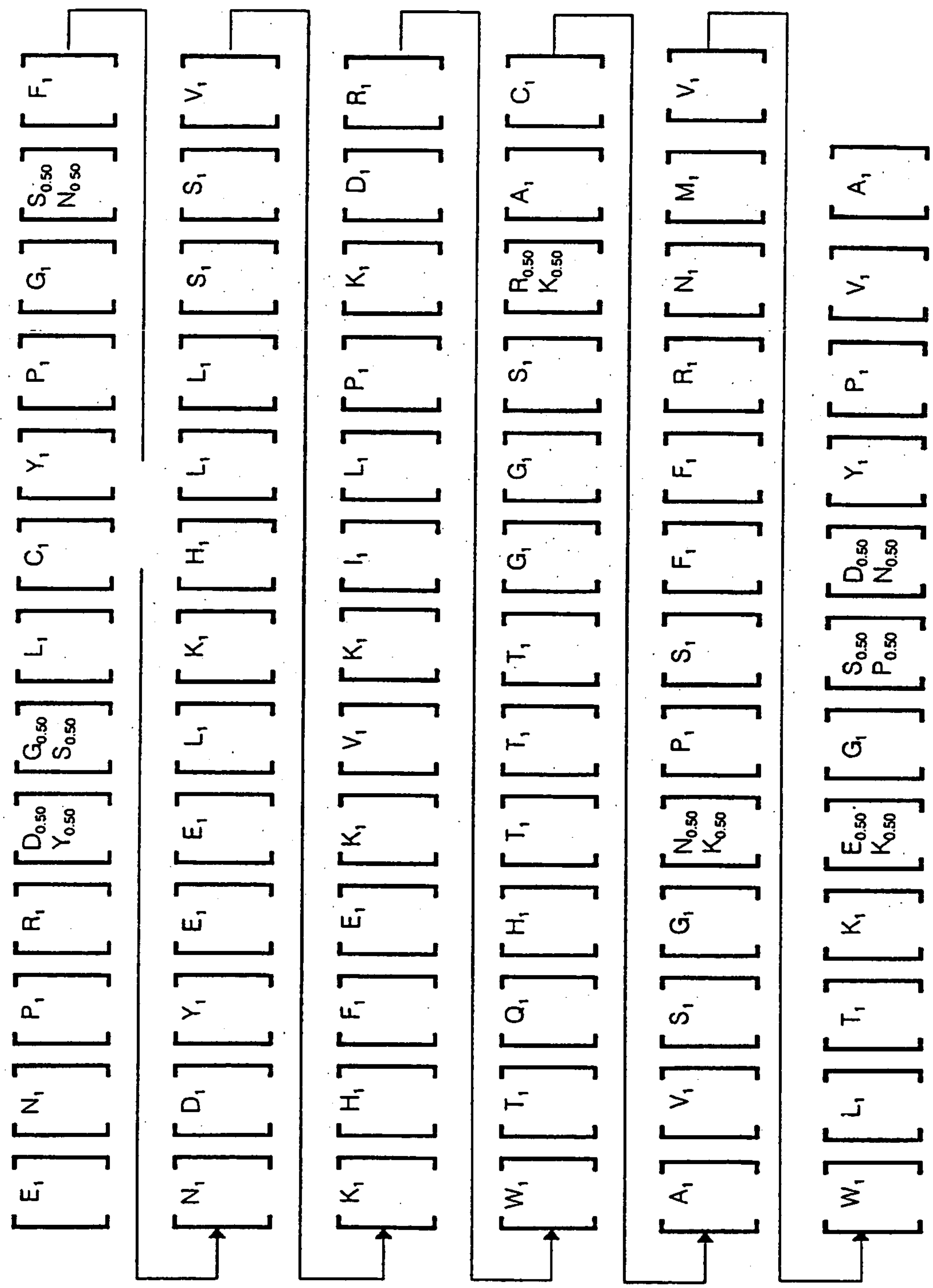

E_1 N_1 P_1 R_1 $D_{0.50}$ $Y_{0.50}$ $G_{0.50}$ $S_{0.50}$ L_1 C_1 Y_1 P_1 G_1 $S_{0.50}$ $N_{0.50}$ F_1
N_1 D_1 Y_1 E_1 E_1 L_1 K_1 H_1 L_1 L_1 S_1 S_1 V_1
K_1 H_1 F_1 E_1 K_1 V_1 K_1 I_1 L_1 P_1 K_1 D_1 R_1
W_1 T_1 Q_1 H_1 T_1 T_1 T_1 G_1 G_1 S_1 $R_{0.50}$ $K_{0.50}$ A_1 C_1
A_1 V_1 S_1 G_1 $N_{0.50}$ $K_{0.50}$ P_1 S_1 F_1 F_1 R_1 N_1 M_1 V_1
W_1 L_1 T_1 K_1 $E_{0.50}$ $K_{0.50}$ G_1 $S_{0.50}$ $P_{0.50}$ $D_{0.50}$ $N_{0.50}$ Y_1 P_1 V_1 A_1

Figure 19

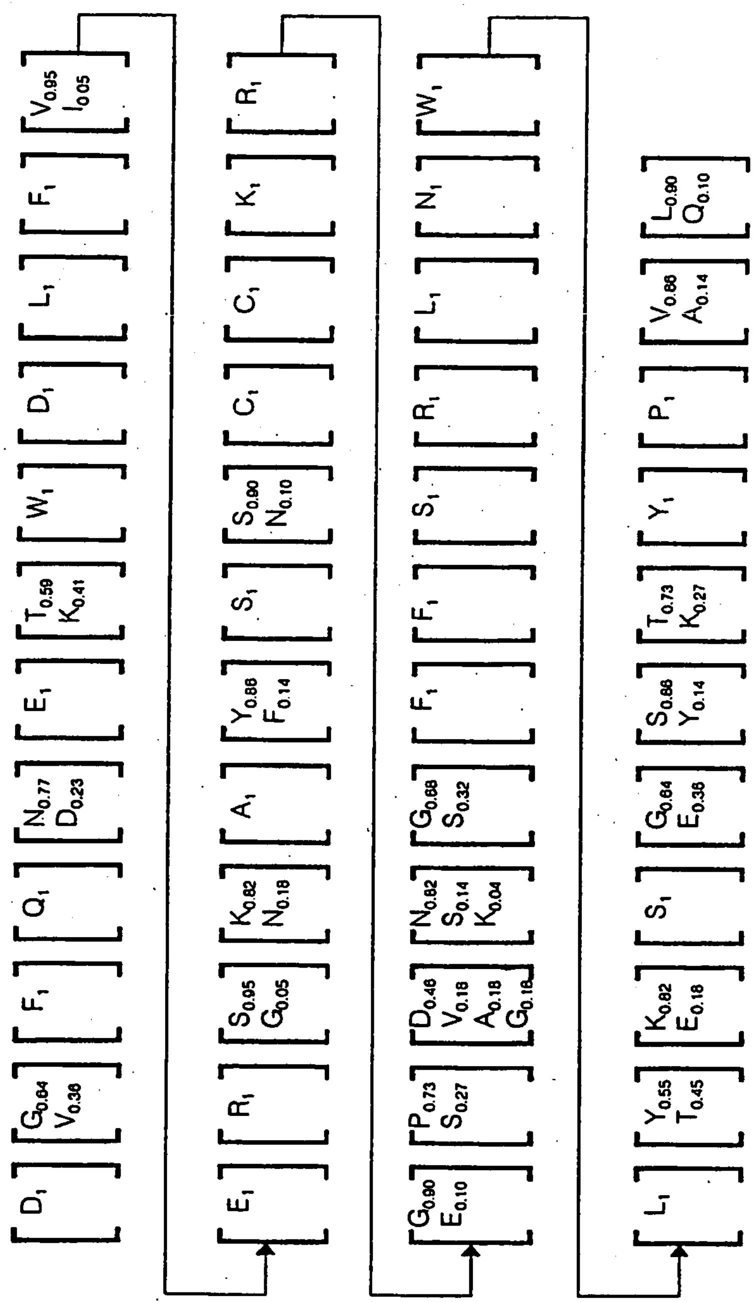

D1
G0.64 V0.36
F1
Q1
N0.77 D0.23
E1
T0.59 K0.41
W1
D1
L1
F1
V0.95 I0.05
E1
R1
S0.95 G0.05
K0.82 N0.18
A1
Y0.86 F0.14
S1
S0.90 N0.10
C1
C1
K1
R1
G0.90 E0.10
P0.73 S0.27
D0.46 V0.18 A0.18 G0.18
N0.82 S0.14 K0.04
G0.68 S0.32
F1
F1
S1
R1
L1
N1
W1
L1
Y0.55 T0.45
K0.82 E0.18
S1
G0.64 E0.36
S0.86 Y0.14
T0.73 K0.27
Y1
P1
V0.86 A0.14
L0.90 Q0.10

Figure 20 A

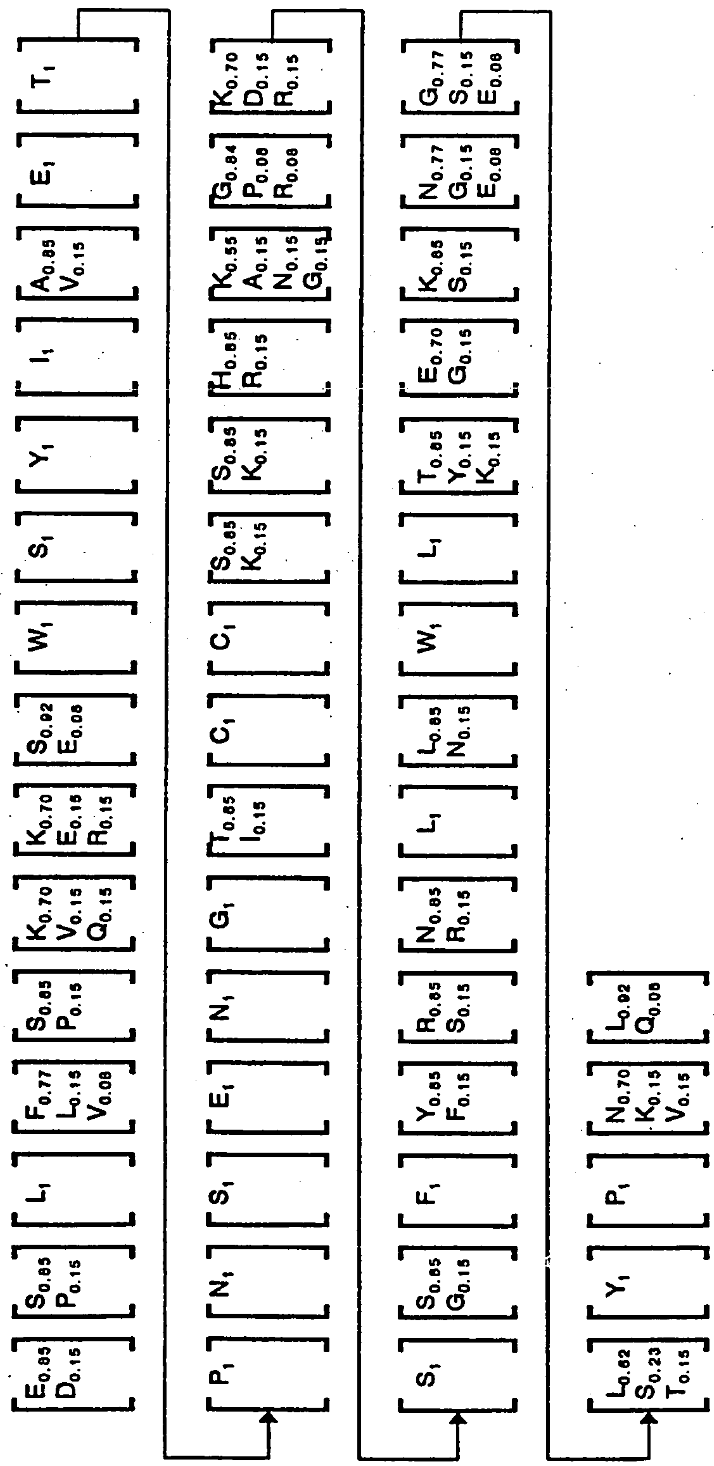


Figure 20 B

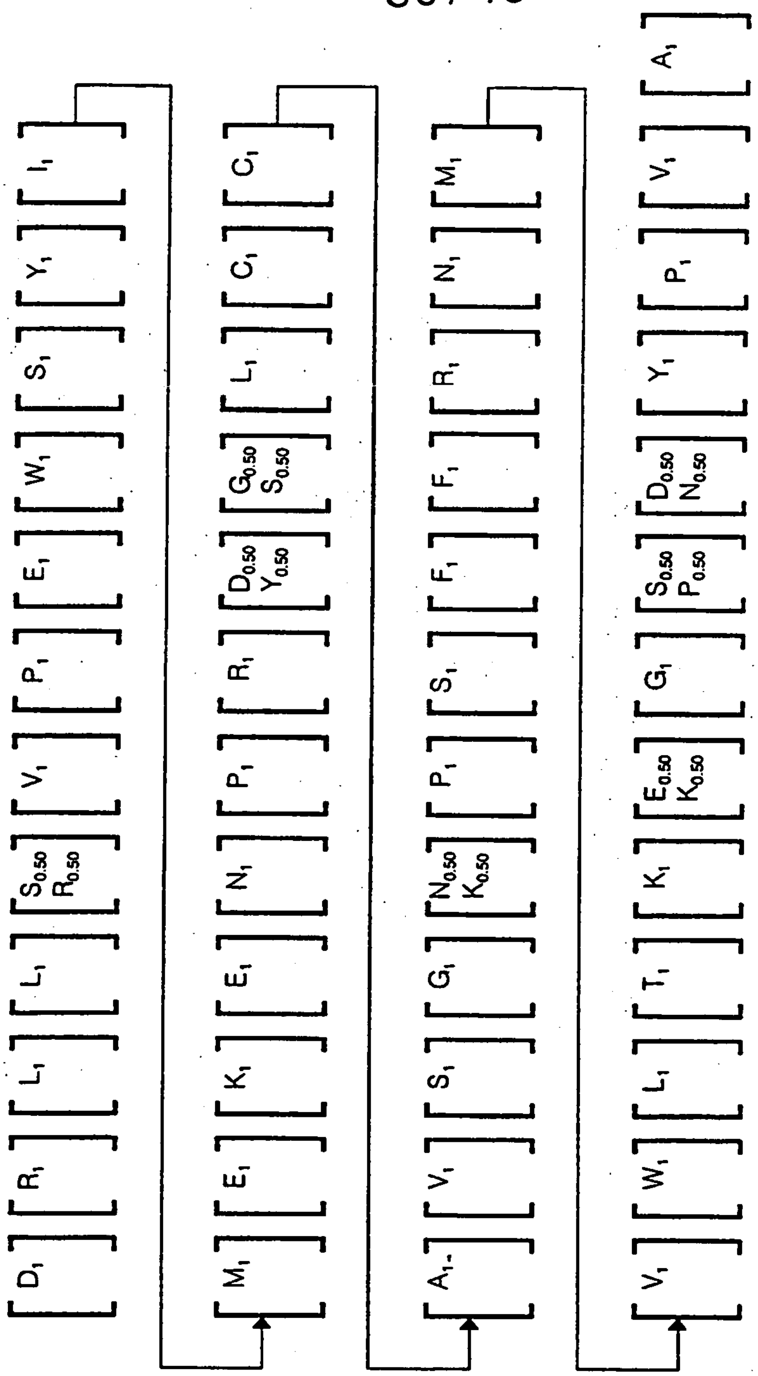

D1 R1 L1 L1 S0.50 R0.50 V1 P1 E1 W1 S1 Y1 I1
M1 E1 K1 E1 N1 P1 R1 D0.50 Y0.50 G0.50 S0.50 L1 C1 C1
A1 V1 S1 G1 N0.50 K0.50 P1 S1 F1 F1 R1 N1 M1
V1 W1 L1 T1 K1 E0.50 K0.50 G1 S0.50 P0.50 D0.50 N0.50 Y1 P1 V1 A1

Figure 20 C

	E_8	F_8	Q_8	M_8	G_8	A_8	A_7	P_8	T_8	T_8	S_5	D_8	V_6	A_7	G_7	L_8	E_3	K_2	D_8	P_8	V_2	A_1	N_8	V_8	A_8	R_8
							E_1				N_1		A_1	E_1	D_1		Q_4	N_6			T_5	T_6				
											K_1		T_1				S_1				K_1	V_1				
											R_1															
A_1	E	F	Q	M	G	A	A	P	T	T	S	D	V	A	G	L	E	K	D	P	V	A	N	V	A	R
A_2	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.	.	.	.	.	.	V	.	.	.	.
C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q	N	.	.	T	T	.	.	.	.
H	.	.	.	.	.	.	.	.	.	.	N	.	A	.	D	.	Q	N	.	.	K	T	.	.	.	.
I	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	N	.	.	T	T	.	.	.	.
J	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Q	N	.	.	T	T	.	.	.	.
K	.	.	.	.	.	.	.	.	.	.	.	.	.	E	.	.	Q	N	.	.	T	T	.	.	.	.
L_3	.	.	.	.	.	.	E	.	.	.	.	.	T	.	.	.	S	N	.	.	T	T	.	.	.	.

Figure 21 A

	E_6	F_6	Q_6	M_6	G_6	A_5	K_6	P_6	T_6	T_3	T_4	T_6	G_6	N_6	A_5	V_1	A_6	P_6	S_5	T_6	L_4	T_6	A_6	R_6
						D_1				A_2	D_1				S_1	A_3			T_1		C_2			
										S_1	S_1					T_2								
B	E	F	Q	M	G	A	K	P	T	T	T	T	G	N	A	V	A	P	S	T	L	T	A	R
Ba	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	T	-	-	-	-	-	-	-	-
D	-	-	-	-	-	-	-	-	-	-	D	-	-	-	S	A	-	-	-	-	-	-	-	-
E	-	-	-	-	-	D	-	-	-	S	-	-	-	-	-	T	-	-	T	-	-	-	-	-
L_1	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-	-	C	-	-	-
L_2	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	A	-	-	-	-	C	-	-	-

Figure 21 B

	L_2	A_3	K_1	P_1	V_2	L_2	D_8	T_4	T_8	T_8	L_8	N_8	P_8	T_8	I_8	A_8	G_8	K_2	G_8	T_2	V_8	V_2	S_2	S_6	A_8	E_8	N_2	E_3	L_8	A_2
	S_6		T_4	A_4	I_6	F_6		V_4										A_6		D_5		K_6	T_3	G_1			G_6	Q_5		G_6
			E_3	T_3																E_1			A_3	N_1						
A_1	L	A	K	P	V	L	D	T	T	T	L	N	P	T	I	A	G	K	G	T	V	V	S	S	A	E	N	E	L	A
I	-	-	E	A	I	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B	S	-	E	T	I	F	-	V	-	-	-	-	-	-	-	-	-	A	-	D	-	K	T	-	-	-	G	Q	-	G
Ba	S	-	E	T	I	F	-	V	-	-	-	-	-	-	-	-	-	A	-	D	-	K	T	-	-	-	G	Q	-	G
D	S	-	T	A	I	F	-	-	-	-	-	-	-	-	-	-	-	A	-	D	-	K	T	G	-	-	G	E	-	G
E	S	-	T	A	I	F	-	-	-	-	-	-	-	-	-	-	-	A	-	D	-	K	A	-	-	-	G	Q	-	G
L1	S	-	T	A	I	F	-	-	-	-	-	-	-	-	-	-	-	A	-	E	-	K	A	N	-	-	G	Q	-	G
L2	S	-	T	T	-	F	-	V	-	-	-	-	-	-	-	-	-	A	-	D	-	K	A	-	-	-	G	Q	-	G

Figure 22 A

	L_7	A_6	E_5	A_6	I_5	L_6	D_7	V_6	T_7	T_7	L_7	N_7	P_7	T_7	I_7	A_6	G_7	K_5	G_7	S_5	V_7	V_6	S_2	A_3	G_5	T_2	D_3	N_5	E_1	L_5	A_5
		V_1	T_1	P_1	V_2	V_1		I_1								T_1		C_2		T_1		A_1	A_4	S_4	N_2	S_5	E_4	G_2	Q_2	I_2	S_2
			K_1																	A_1			G_1						D_4		
C	L	A	E	A	I	L	D	V	T	T	L	N	P	T	I	A	G	K	G	S	V	V	S	A	G	T	D	N	E	L	A
H	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	A	S	-	S	-	-	D	-	-
J	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	S	-	S	E	-	D	-	-
K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	A	-	-	-	S	-	S	-	-	-	-	-
L_3	-	-	-	-	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	S	-	S	E	-	-	-	-
F	-	V	T	P	V	V	-	I	-	-	-	-	-	-	-	-	-	C	-	-	-	A	G	-	N	-	E	G	Q	I	S
G	-	-	K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	A	-	N	S	E	G	Q	I	S

Figure 22 B

$D_{57}S_{57}S_{55}A_{57}H_{57}S_{37}T_{56}P_{57}S_{57}T_{57}P_{57}A_{33}$

T_{2} $G_{20}A_{1}$ V_{24}

	D	S	S	A	H	S	T	P	S	T	P	A
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	G	A	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
NA	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	G	.	.	.	.	.	.

	D	S	S	A	H	S	T	P	S	T	P	A
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	T	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
SAPA	.	.	T	.	.	G	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	.	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.
	.	.	.	.	.	G	.	.	.	.	.	V
	.	.	.	.	.	.	.	.	.	.	.	.

Figure 23

$V_{19}A_{19}E_{21}A_{19}E_{21}K_{21}Q_{21}R_{4}$ $A_{21}A_{21}E_{21}A_{21}T_{15}K_{21}$

A_{2} V_{2} T_{2} K_{17} M_{4}

A_{2}

CRA

V	A	E	A	E	K	Q	R	A	A	E	A	T	K
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	M	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	M	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	A	-
A	V	-	T	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	A	-
A	V	-	T	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	M	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	M	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	K	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 24

```
C33K33K33N31N33K3 Q33K33V18V33Q30P33E33N33L33E33Y33T33I29V33I28T33P33H33S33Q33E33E33H22A33V33G33N33D33T33G33K33H33G33K32E33I27K33I21T33P33Q33S33S33I31T32E33A33E33L33T33G33Y33G33T33V32T33W33E31C33
   Q1 S2 E30  I18 L2  V4 V5  N10  E1 L1 V12  N1 A1  I1 Q1
              F1  H1  V4 V5  S1      V6      T1        Q1

New Guinea C-44   CKKNNKQKVVQPENLEYTIVITPHSQEEHAVGNDTGKHGKEIKITPQSSITEAELTGYGTVTWEC
Sri Lanka-68      ............................................................I....
Sri Lanka-68      .....E...........................................N...............
Philippines-83    .....E..I........................................................
Taiwan-87         .....E..I........................................................
Thailand-58       .................................................................
Thailand-84       .....E...........................................T...............
Thailand-80b      .....E..I...........V....................V.......................
Thailand-80d      .....E..I...........V....................V.......A...............
Thailand-80e      .....E..............V....................V.......................
Malaysia-67a      .Q...E..I...........V..........................................Q.
Malaysia-67b      .....E..I...........V............................................
Thailand-80a      .....E..I........................................................
Thailand-80c      .....E..I..................................V.....................
Jamaica-83        .....E....L......................................................
Brazil-90         .....E....L............................E.........................
Malaysia-88       .....E..F.H..............................L.....................Q.
Indonesia-76      .....E......................N....................................
Somalia-84        ...S.E......................N....................................
Burkina Faso-83   ...S.E......................N....................................
Seychelles-77a    .....E......................N..............V.....................
Seychelles-77b    .....E..I...................N..............V.....................
Sri Lanka-82      .....E......................S..............V.....................
Sri Lanka-85a     .....E......................N..............V.....................
Sri Lanka-85b     .....E......................N..............V.....................
Sri Lanka-89      .....E..I...................N..............V.....................
Sri Lanka-90b     .....E..I...................N..............V.....................
Sri Lanka-90a     .....E......................N..............V.....................
Trinidad-83       ........I................................V.......................
India-57          .....E..I.........V......................V.......................
Puerto Rico-69    .....E..I.........V......................V.......................
Puerto Rico 81    .....E..I.........V......................V.......................
Tonga-74          .....E..I.........V......................V.......................
```

Figure 25 A

	I_{32}	T_{33}	V_{33}	N_{32}	P_{33}	I_{33}	V_{31}	T_{33}	E_{33}	K_{33}	D_{33}	S_{32}	P_{33}	V_{33}	N_{32}	I_{31}	E_{32}
	T_1			Y_1			A_1					R_1			K_1	V_2	D_1
							G_1										
New Guinea C-44	I	T	V	N	P	I	V	T	E	K	D	S	P	V	N	I	E
Sri Lanka-68	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Philippines-83	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Taiwan-87	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thailand-58	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thailand-64	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thailand-80b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	K	.	.
Thailand-80d	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thailand-80e	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Malaysia-87a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.
Malaysia-87b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.
Thailand-80a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Thailand-80c	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Jamaica-83	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Brazil-90	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Malaysia-86	.	.	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	D
Indonesia-76	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Somalia-84	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Burkina Faso-83	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Seychelles-77a	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
Seychelles-77b	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.
Sri Lanka-82	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-85a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-85b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-89	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-90b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Sri Lanka-90a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Trinidad-53	.	.	.	.	.	.	.	.	.	.	.	R	.	.	.	.	.
India-57	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Puerto Rico-69	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Puerto Rico 81	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Tonga-74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 25 B

1 2 25 28 32 35 42

$T_1V_1T_1R_2G_2T_2I_1S_1D_1P_2Q_1R_1A_2K_1E_2A_1L_2D_1K_2Y_2E_1L_1E_2N_3H_1D_1L_3K_1T_1K_1N_3E_1G_1L_3K_2T_2E_2N_2E_2G_1L_2$

$N_1G_1D_1G_1N_1P_1R_1E_1V_1I_1E_1D_1L_1L_1A_1N_2N_1P_1A_1I_1Q_1$ I_1R_1 $R_1H_1E_1$ K_1D_1 $T_1A_1R_1L_1N_1N_1A_1$

$R_1V_1F_1P_1$ $V_1E_1N_1$ D_1K_1 R_1 D_1V_1 S_1M_1 $Q_1A_1N_1$ D_1 K_1 L_1

TYPE	
5	T V T R G T I S D P Q R A K E A L D K Y E L E N H D L K T K N E G L K T E N E G L
1	N G D G N P R E V I E D L A A N N P A I Q - I R - R H E - K D - K A R L E N A
6	R V F P - - - V E N - D K - R - L - N - - D V - - S M - Q A N - D K - T - - - N L -
19	V R Y S R E S P E D K - - K I I D D - D A K E N E -
24	V H T - S Q T D T S E K V Q E R A D S F E I E - N T - - L

Figure 26

P_4	E_3	G_4	A_3	Q_3	G_4	P_3	R_4	G_4	E_3	P_3	G_4
	I_1		P_1	P_1		S_1			R_1	S_1	

P	E	G	A	Q	G	P	R	G	E	P	G
-	-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	S	-	-	-	-	-
-	I	-	P	P	-	-	-	-	R	S	-

$$\begin{bmatrix} P_1 \end{bmatrix} \begin{bmatrix} E_{0.75} \\ I_{0.25} \end{bmatrix} \begin{bmatrix} G_1 \end{bmatrix} \begin{bmatrix} A_{0.75} \\ P_{0.25} \end{bmatrix} \begin{bmatrix} Q_{0.75} \\ P_{0.25} \end{bmatrix} \begin{bmatrix} G_1 \end{bmatrix} \begin{bmatrix} P_{0.75} \\ S_{0.25} \end{bmatrix} \begin{bmatrix} R_1 \end{bmatrix} \begin{bmatrix} G_1 \end{bmatrix} \begin{bmatrix} E_{0.75} \\ R_{0.25} \end{bmatrix} \begin{bmatrix} P_{0.75} \\ S_{0.25} \end{bmatrix} \begin{bmatrix} G_1 \end{bmatrix}$$

Figure 27

	K_3	E_1	E_3	I_3	E_2	Y_3	L_3	N_2	S_1	I_1	L_3	S_1	Y_2	S_3
		A_1			D_1			E_1	K_2	T_1		P_2	F_1	
		N_1								V_1				
(E-selectin)	K	E	E	I	E	Y	L	N	S	I	L	S	Y	S
(L-selectin)	-	A	-	-	-	-	-	E	K	T	-	P	F	-
(P-selectin)	-	N	-	-	D	-	-	-	K	V	-	P	-	-

$$\begin{bmatrix} K_1 \end{bmatrix} \begin{bmatrix} E_{0.33} \\ A_{0.33} \\ N_{0.34} \end{bmatrix} \begin{bmatrix} E_1 \end{bmatrix} \begin{bmatrix} I_1 \end{bmatrix} \begin{bmatrix} E_{0.67} \\ D_{0.33} \end{bmatrix} \begin{bmatrix} Y_1 \end{bmatrix} \begin{bmatrix} L_1 \end{bmatrix} \begin{bmatrix} N_{0.67} \\ E_{0.33} \end{bmatrix} \begin{bmatrix} S_{0.33} \\ K_{0.67} \end{bmatrix} \begin{bmatrix} I_{0.33} \\ T_{0.33} \\ V_{0.34} \end{bmatrix} \begin{bmatrix} L_1 \end{bmatrix} \begin{bmatrix} S_{0.33} \\ P_{0.67} \end{bmatrix} \begin{bmatrix} Y_{0.67} \\ F_{0.33} \end{bmatrix} \begin{bmatrix} S_1 \end{bmatrix}$$

Figure 28

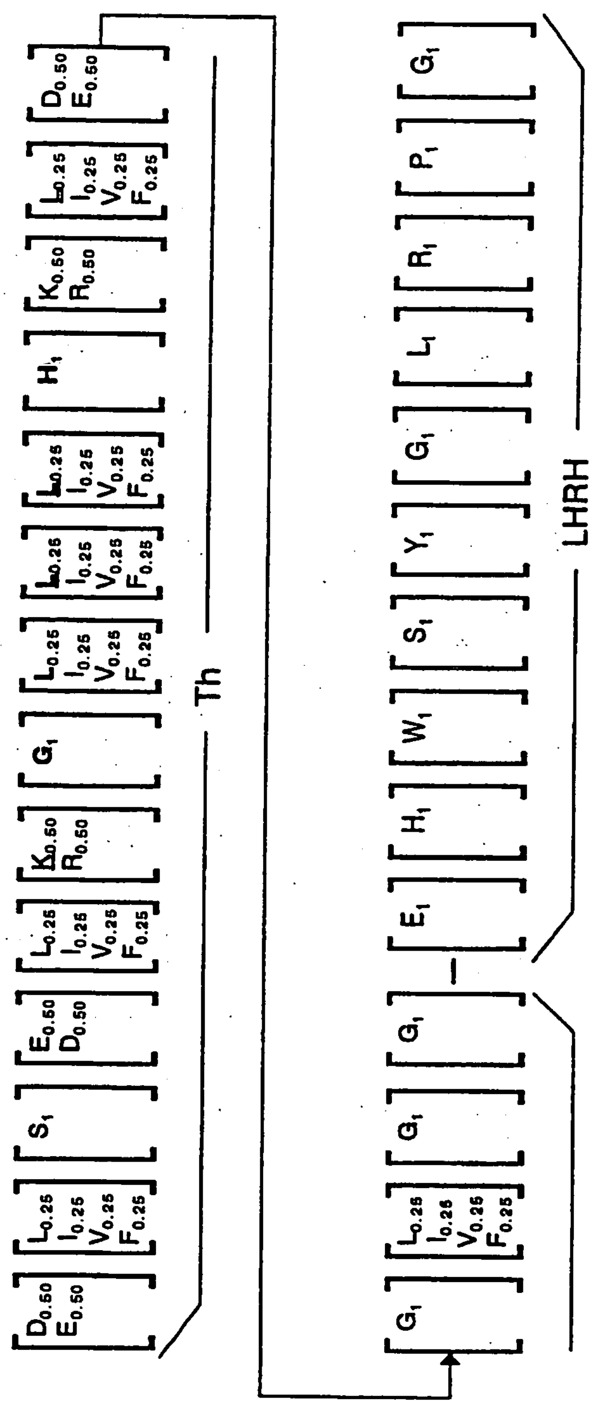


Figure 29 A

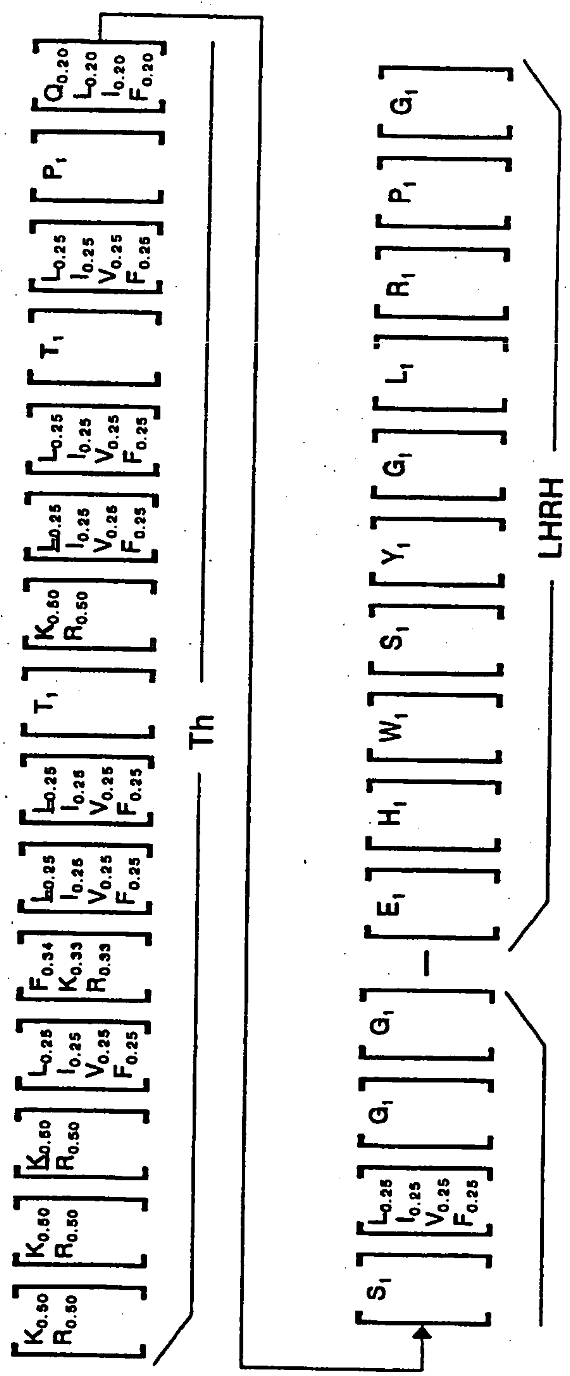


Figure 29 B

R_5	R_7	Y_2	Q_1	K_2	S_1	T_2	E_1	L_1
G_1		I_2	K_3	E_3	I_2	V_2	K_2	K_3
F_1		V_1	L_1	P_1	V_1	I_2	D_1	R_2
		W_1	D_1	G_1	A_1	G_1	V_1	A_1
		S_1	N_1		P_1		L_1	
					L_1		H_1	

HISTONEH3	R	R	Y	Q	K	S	T	E	L
Hsp89α	-	-	I	K	E	I	V	K	K
hSP89β	-	-	V	K	E	V	V	K	K
HEF2	-	-	W	L	P	A	G	D	A
Helicase	-	-	S	K	E	I	-	V	R
Ribosomal protein	G	-	I	D	-	P	I	L	K
L28	F	-	-	N	G	L	I	H	R

$R_{0.71}$	R_1	$Y_{0.28}$	$Q_{0.15}$	$K_{0.28}$	$S_{0.15}$	$T_{0.28}$	$E_{0.15}$	$L_{0.15}$
$G_{0.14}$		$I_{0.28}$	$K_{0.42}$	$E_{0.42}$	$I_{0.28}$	$V_{0.28}$	$K_{0.28}$	$K_{0.42}$
$F_{0.15}$		$V_{0.15}$	$L_{0.15}$	$P_{0.15}$	$V_{0.15}$	$I_{0.28}$	$D_{0.15}$	$R_{0.28}$
		$W_{0.15}$	$D_{0.14}$	$G_{0.15}$	$A_{0.14}$	$G_{0.15}$	$V_{0.14}$	$A_{0.15}$
		$S_{0.14}$	$N_{0.14}$		$P_{0.14}$		$L_{0.14}$	
					$L_{0.14}$		$H_{0.14}$	

Figure 30

	K_8	R_9	W_8	I_8	I_5	L_5	G_8	L_9	N_8	K_8
	R_1		V_1	V_1	V_3	I_2	R_1		Q_1	R_1
					Q_1	M_2				
HIV-1 MN	K	R	W	I	I	L	G	L	N	K
HIV-1 ELI	-	-	-	-	V	-	-	-	-	-
HIV-2 ROD	R	-	V	-	Q	I	-	-	Q	-
HIV PATIENT	-	-	-	-	-	M	-	-	-	-
HIV PATIENT	-	-	-	V	-	-	-	-	-	-
HIV PATIENT	-	-	-	-	-	-	-	-	-	R
HIV PATIENT	-	-	-	-	-	I	R	-	-	-
HIV PATIENT	-	-	-	-	V	-	-	-	-	-
HIV PATIENT	-	-	-	-	V	M	-	-	-	-

$$\begin{bmatrix} K_{0.89} \\ R_{0.11} \end{bmatrix} \begin{bmatrix} R_1 \\ \ \end{bmatrix} \begin{bmatrix} W_{0.89} \\ V_{0.11} \end{bmatrix} \begin{bmatrix} I_{0.89} \\ V_{0.11} \end{bmatrix} \begin{bmatrix} I_{0.55} \\ V_{0.33} \\ Q_{0.12} \end{bmatrix} \begin{bmatrix} L_{0.55} \\ I_{0.23} \\ M_{0.22} \end{bmatrix} \begin{bmatrix} G_{0.89} \\ R_{0.11} \end{bmatrix} \begin{bmatrix} L_1 \\ \ \end{bmatrix} \begin{bmatrix} N_{0.89} \\ Q_{0.11} \end{bmatrix} \begin{bmatrix} K_{0.89} \\ R_{0.11} \end{bmatrix}$$

Figure 31

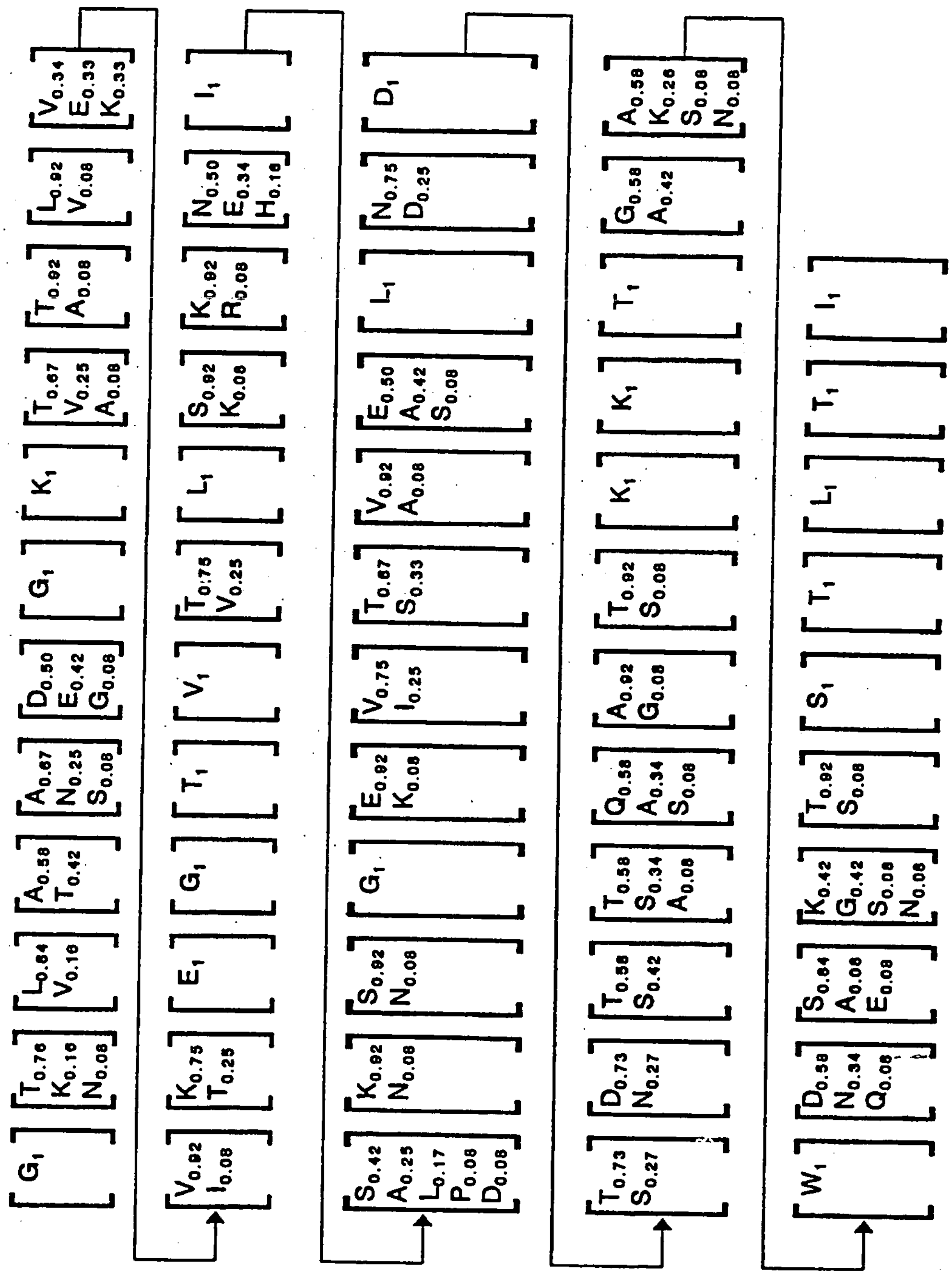


Figure 32 A

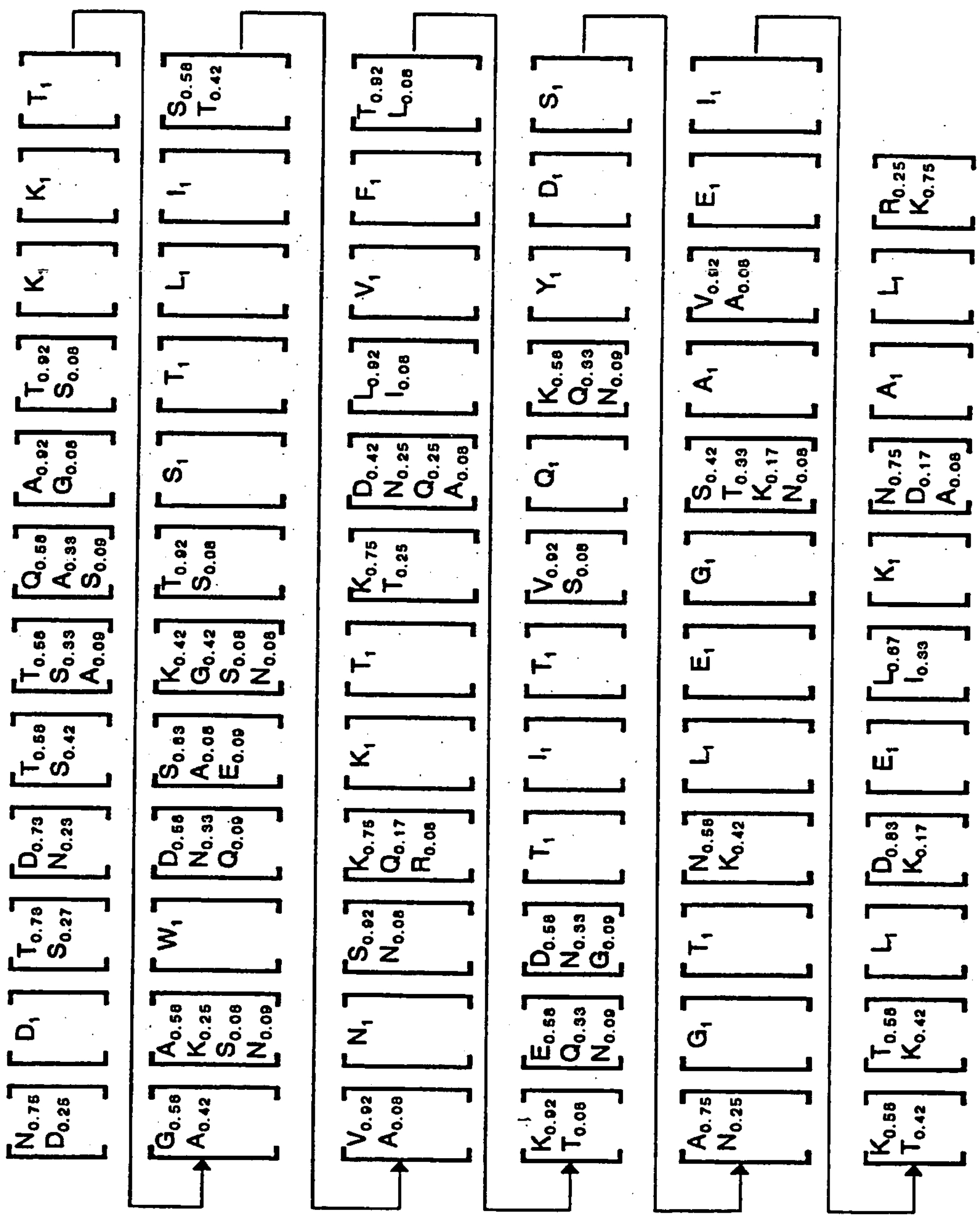
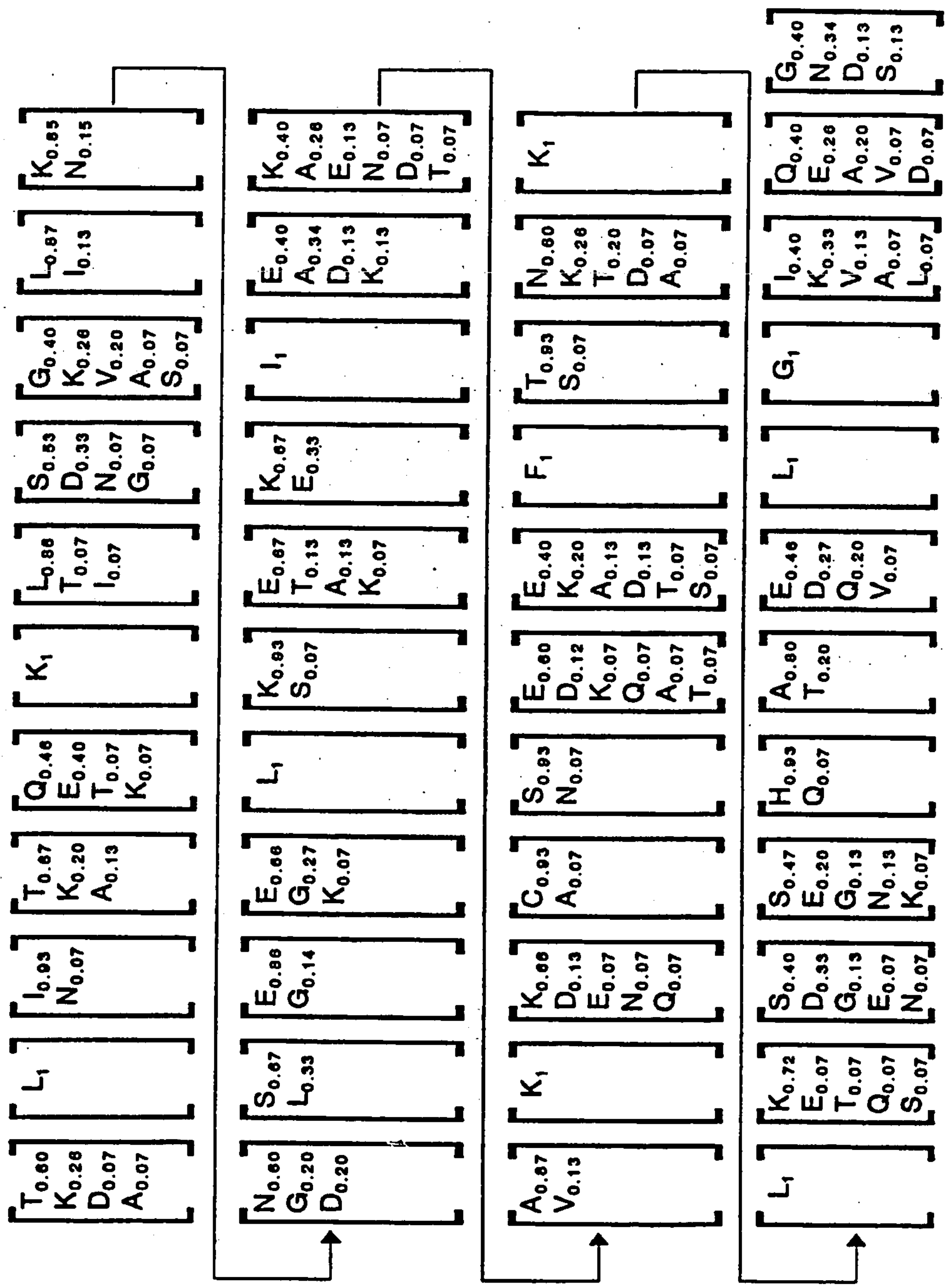


Figure 32 B



$T_{0.60}$ $K_{0.26}$ $D_{0.07}$ $A_{0.07}$
L_1
$I_{0.93}$ $N_{0.07}$
$T_{0.67}$ $K_{0.20}$ $A_{0.13}$
$Q_{0.46}$ $E_{0.40}$ $T_{0.07}$ $K_{0.07}$
K_1
$L_{0.86}$ $T_{0.07}$ $I_{0.07}$
$S_{0.53}$ $D_{0.33}$ $N_{0.07}$ $G_{0.07}$
$G_{0.40}$ $K_{0.26}$ $V_{0.20}$ $A_{0.07}$ $S_{0.07}$
$L_{0.87}$ $I_{0.13}$
$K_{0.85}$ $N_{0.15}$
$N_{0.60}$ $G_{0.20}$ $D_{0.20}$
$S_{0.67}$ $L_{0.33}$
$E_{0.86}$ $G_{0.14}$
$E_{0.66}$ $G_{0.27}$ $K_{0.07}$
L_1
$K_{0.93}$ $S_{0.07}$
$E_{0.67}$ $T_{0.13}$ $A_{0.13}$ $K_{0.07}$
$K_{0.67}$ $E_{0.33}$
I_1
$E_{0.40}$ $A_{0.34}$ $D_{0.13}$ $K_{0.13}$
$K_{0.40}$ $A_{0.26}$ $E_{0.13}$ $N_{0.07}$ $D_{0.07}$ $T_{0.07}$
$A_{0.87}$ $V_{0.13}$
K_1
$K_{0.66}$ $D_{0.13}$ $E_{0.07}$ $N_{0.07}$ $Q_{0.07}$
$C_{0.93}$ $A_{0.07}$
$S_{0.93}$ $N_{0.07}$
$E_{0.60}$ $D_{0.12}$ $K_{0.07}$ $Q_{0.07}$ $A_{0.07}$ $T_{0.07}$
$E_{0.40}$ $K_{0.20}$ $A_{0.13}$ $D_{0.13}$ $T_{0.07}$ $S_{0.07}$
F_1
$T_{0.93}$ $S_{0.07}$
$N_{0.60}$ $K_{0.26}$ $T_{0.20}$ $D_{0.07}$ $A_{0.07}$
K_1
L_1
$K_{0.72}$ $E_{0.07}$ $T_{0.07}$ $Q_{0.07}$ $S_{0.07}$
$S_{0.40}$ $D_{0.33}$ $G_{0.13}$ $E_{0.07}$ $N_{0.07}$
$S_{0.47}$ $E_{0.20}$ $G_{0.13}$ $N_{0.13}$ $K_{0.07}$
$H_{0.93}$ $Q_{0.07}$
$A_{0.80}$ $T_{0.20}$
$E_{0.46}$ $D_{0.27}$ $Q_{0.20}$ $V_{0.07}$
L_1
G_1
$I_{0.40}$ $K_{0.33}$ $V_{0.13}$ $A_{0.07}$ $L_{0.07}$
$Q_{0.40}$ $E_{0.26}$ $A_{0.20}$ $V_{0.07}$ $D_{0.07}$
$G_{0.40}$ $N_{0.34}$ $D_{0.13}$ $S_{0.13}$

Figure 33

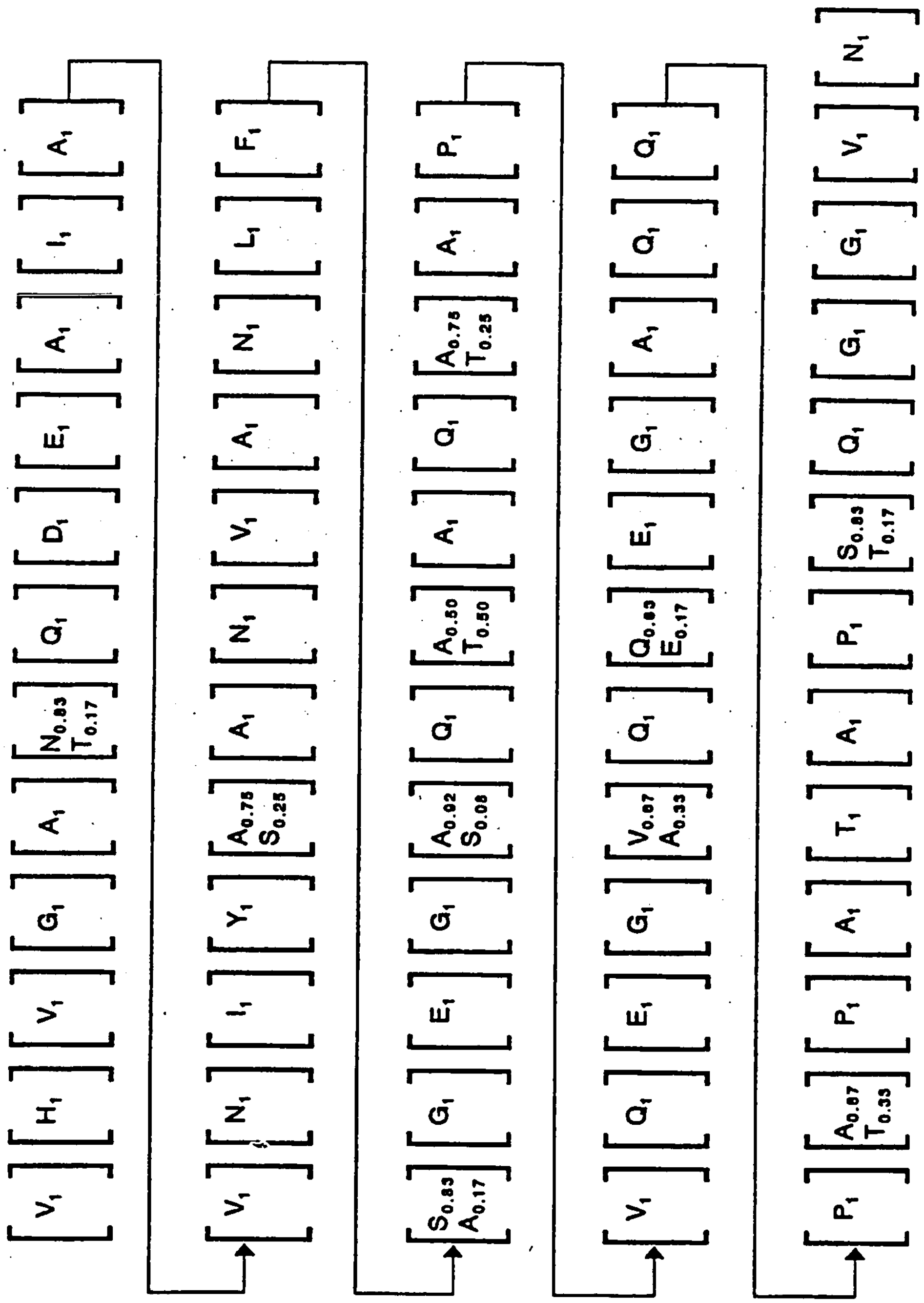

V_1 H_1 V_1 G_1 A_1 $N_{0.83}T_{0.17}$ Q_1 D_1 E_1 A_1 I_1 A_1
V_1 N_1 I_1 Y_1 $A_{0.75}S_{0.25}$ A_1 N_1 V_1 A_1 N_1 L_1 F_1
$S_{0.83}A_{0.17}$ G_1 E_1 G_1 $A_{0.92}S_{0.08}$ Q_1 $A_{0.50}T_{0.50}$ A_1 Q_1 $A_{0.75}T_{0.25}$ A_1 P_1
V_1 Q_1 E_1 G_1 $V_{0.67}A_{0.33}$ Q_1 $Q_{0.83}E_{0.17}$ E_1 G_1 A_1 Q_1 Q_1
P_1 $A_{0.67}T_{0.33}$ P_1 A_1 T_1 A_1 P_1 $S_{0.83}T_{0.17}$ Q_1 G_1 G_1 V_1 N_1

Figure 34

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/12268

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12Q 1/70; A61K 39/29; C07K 14/18
US CL :435/5; 424/228.1; 530/350
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/7.25, 7.92

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, STIC sequence search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,106,726 (WANG) 21 April 1992, especially col. 15, lines 5-9 and 18-23	1-12, 14, 16-19, 21
Y	Proceedings of the National Academy of Sciences USA, Volume 89, issued 15 April 1992, W. Ching et al., "Interaction of immune sera with synthetic peptides corresponding to the structural protein region of hepatitis C virus," pages 3190-3194, especially the Abstract and page 3190, first paragraph.	1-12, 14, 16-19, 21

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
24 JANUARY 1995	07 FEB 1995

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	DONNA C. WORTMAN D. Vizza for
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/12268

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-12, 14, 16-19, 21, the first appearing specie.

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.
Group I, claim(s) 1-12, 14, 16-21, drawn to a peptide composition, a method of making a peptide composition, and a method of using a peptide composition.
Group II, claim(s) 13 and 15, drawn to a method of using a peptide composition.
The inventions listed as Groups I and II do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The claims of group II are drawn to a second method of using a peptide composition which requires distinctly different method steps and reagents and has different goals and outcomes than the method of use of Group I. PCT Rule 13.1 does not provide for a second method of use.
Additionally, Group I contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be examined, the appropriate additional examination fees must be paid.
The species are as follows:
The binding peptides represented by each of SEQ ID NO:1-43.The claims are deemed to correspond to the species listed above in the following manner:
Claims 1-18 and 21 correspond to all the species, and are therefore generic.
Claim 19 corresponds to the species of SEQ ID NO:1-31 and 37-43.
Claim 20 corresponds to the species of SEQ ID NO: 32-36

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: Each binding peptide represented by a different sequence does not have the same or corresponding special technical feature as each other binding peptide since each recited peptide is a distinct chemical entity with binding properties and variability which differ from each of the other peptides. Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.